DOCUMENT RESUME

ED 045 327 RE 003 183

AUTHOR Ho, Wai-Ching: And Others
TITLE Longitudinal Effects of i.t.a. on Pupil's Reading Achievement Through Grade Three.
INSTITUTION Educational Research Council of America, Cleveland, Ohio.
PUB DATE Mar 70
NOTE 43p.

EDRS PRICE EDRS Price MF-$0.25 HC-$2.25
DESCRIPTORS Basic Reading, *Initial Teaching Alphabet, *Kindergarten, *Longitudinal Studies, Phonics, *Primary Grades, Reading Achievement, *Reading Research, Spelling

ABSTRACT

To make this report more readable for teachers and still present all the data, it was compiled in two parts: (1) Report to Teachers and (2) Technical Report. The investigation was conducted (1) to compare the longitudinal effects of the initial teaching alphabet (i.t.a.) and traditional orthography (T.O.) beginning reading programs through and end of grade 3 and (2) to determine the feasibility of starting i.t.a. instruction in kindergarten. Approximately 700 pupils from nine school districts were grouped so that one-third of the pupils started T.O. in grade 1, one-third started i.t.a. in grade 1, and the remaining one-third started i.t.a. in kindergarten. The i.t.a. groups transferred to T.O. basal readers during the second semester of grade 2. Tests in i.t.a. in grade 1 demonstrated the superiority of i.t.a. pupils over the T.O. pupils in skills requiring sound-letter association. After the transition from i.t.a. to T.O., few significant differences were found between the groups on spelling and other reading subtests. Since the group with which i.t.a. instruction was begun in kindergarten maintained their advantage through the end of grade 3, it was suggested that it seems not only feasible but necessary to introduce i.t.a. in kindergarten. Tables and references are included. (DH)

EDO 45327

LONGITUDINAL EFFECTS OF i.t.a. ON PUPILS' READING ACHIEVEMENT THROUGH GRADE THREE

PREPARED BY

Wai-Ching Ho

Charles F. Eiszler Vickie Stroh

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
Educational Research Council of America
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE US OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER"

RE 003 183

Evaluation and Testing Department

EDUCATIONAL RESEARCH COUNCIL OF AMERICA

March 1970

Participating School Districts*

SCHOOL DISTRICTS	CHIEF EXECUTIVES
Aurora Public Schools	Mr. H. Paul Snyder
Avon Lake Public Schools	Mr. Robert J. Robinson
Bay Village	Dr. Robert C. Cawrse
Berea	Dr. Frank Mayer
Board of Catholic Education (Cleveland Diocese)	The Very Rev. Msgr. William N. Novicky
Bradford Area Schools (Bradford, Pennsylvania)	Mr. Frederick Shuey
Brockton Public Schools (Brockton, Massachusetts)	Dr. Anthony D'Antuono
Brooklyn	Dr. Keith Wallace
Chardon	Dr. Arthur R. Williamson
Cuyahoga Heights	Dr. John R. Thorson
Fairview Park	Dr. Frank W. Barr
Greenville Public Schools Greenville, Michigan	Mr. Burl A. Glendening
Independence	Mr. Stanley E. Skoczen
Lakewood	Dr. John Ellis
Lutheran Schools (Greater Cleveland)	Dr. E. F. Sagehorn
Mayfield	Mr. Joseph L. Baird
Muskegon Public Schools (Muskegon, Michigan)	Dr. William L. Austin
Niles Public Schools (Niles, Michigan)	Mr. Richard B. Warren
North Olmsted	Dr. Robert Van Auken
Olmsted Falls	Dr. Donald Cobb
Owatonna Public Schools (Owatonna, Minnesota)	Mr. C.P. Mickelson
Summit Public Schools (Summit, New Jersey)	Dr. Robert Salisbury
Westlake	Dr. Kenneth W. Harris
Wickliffe	Mr. Paul J. Wallace

* March 1970

ED0 45327

LONGITUDINAL EFFECTS OF i.t.a. ON PUPILS' READING ACHIEVEMENT THROUGH GRADE THREE

PREPARED BY

Wai-Ching Ho, Charles F. Eiszler, Vickie Stroh

Part I Report to Teachers

Part 2 Technical Report

Appendix A Hierarchal Analysis of Variance with Unequal Sample Sizes (Maurice M. Tatsuoka)

Appendix B Mean Score Profiles for Grades One, Two and Three

The research reported herein was a cooperative research project of the Educational Research Council of America and its member schools.

EDUCATIONAL RESEARCH COUNCIL OF AMERICA / CLEVELAND, OHIO

ERC SCHOOL DISTRICTS THAT CONTRIBUTED TO THE STUDY

Aurora, Ohio
Berea, Ohio
Board of Catholic Education
(Cleveland Diocese)
Brockton, Massachusetts
Chardon, Ohio
Lutheran Schools
(Greater Cleveland)
Mayfield, Ohio
Muskegon, Michigan
Owatonna, Minnesota

CONSULTANTS TO THE STUDY

Dr. Doyle Bishop	University of Illinois
Dr. John B. Carroll	Educational Testing Service
Dr. J. Thomas Hastings	University of Illinois
Dr. Maurice Tatsuoka	University of Illinois

Printed at the Offices of
Educational Research Council of America
Rockefeller Bldg. — Cleveland, Ohio 44113

Copyright © 1970 Educational Research Council of America
All rights reserved. Printed in the United States of America.

TABLE OF CONTENTS

Preface v

Part 1 — Report to Teachers 1

Part 2 — Technical Report 4

Appendix A — Hierarchal Analysis of Variance with Unequal Sample Sizes 21

Appendix B — Mean Score Profiles for Grades One, Two, and Three 29

References 35

PREFACE

This report is an attempt to combine the report to teachers and the technical details in one volume. To make the report to teachers in Part 1 more readable, only the essential procedure and findings are reported. Detailed discussion of the technical aspects and supporting data are given in Part 2. Appendix A, written by Dr. Maurice Tatsuoka, presents a full explanation of the hierarchal analysis of variance, the technique used in the present study. The profiles in Appendix B are designed to give the readers an idea of the average achievement of various treatment groups at a glance.

The study reported here is a cooperative research venture by the ERC (Educational Research Council of America) and its member school districts. Two major topics were investigated in this study: (1) comparison of the longitudinal effects of i.t.a. (initial teaching alphabet) and T.O. (traditional orthography) beginning reading programs through the end of grade three; (2) feasibility of starting i.t.a. instruction in kindergarten.

The study was undertaken in the school year 1965-66 through 1967-68. Originally, 21 school districts participated in the study. Because of conflicts with the local testing programs of some school districts, change in the membership of ERC school districts, etc. only the nine school districts listed on page ii remained throughout the whole study. When the study was launched, all nine school districts had implemented i.t.a. for two years.

The completion of the study required much cooperation and support of the nine contributing school districts. Acknowledgment is due to the superintendents, i.t.a. coordinators, principals, first, second, and third grade teachers, and other personnel of these schools.

Dr. John B. Carroll and Dr. J. Thomas Hastings were advisers to this study. Dr. Maurice Tatsuoka advised in the planning of the statistical analysis. Dr. Doyle Bishop wrote the computer program and did the actual computation at the University of Illinois.

Special credit should be given to Dr. Thomas Bibler who supervised the production, to Mrs. Susan Detienne who did much to assist in editing, and to the secretarial staff of the Evaluation and Testing Department, ERC, who typed and proofread the report. The Art Department staff of ERC was responsible for the art work.

April, 1970

Wai-Ching Ho
Educational Research Council of America

PART 1

REPORT TO TEACHERS

The importance of success in beginning reading has long been recognized by educators. However, the best method of teaching beginning reading continues to be a controversial subject. To help beginners overcome the stumbling block caused by the inconsistency in symbol-sound relationships in the English language, Sir James Pitman of England devised the Initial Teaching Alphabet (i.t.a.). It consists of 44 characters designed to make the written symbols and the sounds of our language correspond more consistently. Further simplification is achieved through the elimination of capitals. After the pupils gain some fluency in reading i.t.a., transition is made to Traditional Orthography (T.O.), the conventional alphabet and spelling.

With the cooperation of the Council schools, the Evaluation and Testing Department of ERC conducted a longitudinal study to investigate the effects of i.t.a. on the pupils' reading achievement and the feasibility of starting i.t.a. in kindergarten. This study was designed to follow the same pupils through the end of grade three.

The study included approximately 700 pupils from nine Council school districts. About one-third of the pupils started T.O. in grade one, another third started i.t.a. in grade one, and the remaining third started i.t.a. in kindergarten. The T.O. group used various basal readers, predominantly the Ginn and the Scott, Foresman series, supplemented by other readers such as those published by Lippincott. The i.t.a. groups used the Downing and/or the Early To Read series for the initial i.t.a. instruction and generally transferred to T.O. basal readers during the second semester in grade two. The T.O. readers most often used by the i.t.a. groups included the Ginn; Scott, Foresman; and Lippincott series, or some combination of these.

Reading achievement of the pupils in this study was measured by the reading subtests in the Stanford Achievement Test: Word Reading, Paragraph Meaning, Vocabulary, Spelling, and Word Study Skills for grade one; Word Meaning, Paragraph Meaning, Spelling, Word Study Skills, and Language for grades two and three. In the fifth month of grade one, the i.t.a. pupils took the subtests in i.t.a. At the end of grades one, two, and three, all pupils were tested in T.O.

The three groups were compared on each subtest by ability level. Ability levels were determined by the Lorge-Thorndike Intelligence Tests given in grade one. The average IQ was 95 for the low group (ranging from 64 to 103); 108 for the middle group (ranging from 104 to 114); and 121 for the high group (ranging from 115 to 139). Significance of the differences among the three groups was tested by a statistical technique called the nested factor

design (a two-way hierarchal analysis of variance). This method was employed in order to isolate the unique factors that were associated with each classroom. A more detailed description of the technique is found in Appendix A.

The following are some highlights of the results of this study:

- In the fifth month of grade one, when the i.t.a. pupils were tested in i.t.a., they demonstrated superiority to the T.O. pupils in Word Reading, Paragraph Meaning, Spelling, and Word Study Skills at <u>all</u> ability levels.

- At the end of grade one, when the subtests were given in T.O., no significant differences were found between the T.O. and i.t.a. pupils in Word Reading, Paragraph Meaning, Vocabulary, and Word Study Skills at <u>all</u> ability levels, even though over 90% of the i.t.a. pupils had not made the formal transition to T.O.

- At the end of grades two and three, no significant differences were found between the i.t.a. pupils and T.O. pupils on Word Meaning, Paragraph Meaning, Word Study Skills, and Language at <u>all</u> ability levels.

- Despite no statistical significance between the i.t.a. pupils and the T.O. pupils in most subtests beyond grade one, the group that started i.t.a. in kindergarten performed better on almost all the subtests throughout the three years at <u>all</u> ability levels.

- The T.O. reading series used by the i.t.a. pupils in the post-transition period seems crucial to their success in T.O. reading. Those i.t.a. pupils who used a reading series that emphasized the "phonics approach" (either exclusively or combined with a reading series that emphasized the "meaning approach") tended to be superior to their counterparts who used readers which emphasized the "meaning approach." Most pupils of the low ability level who were instructed under a "phonics approach" achieved at or above their grade placement on all subtests at the end of grades two and three.

- Spelling did not seem to cause particular difficulty for i.t.a. pupils of <u>all</u> ability levels after they transferred to T.O. Although at the end of grade two middle ability T.O. pupils scored significantly higher on the Spelling subtest than their i.t.a. counterparts, by the end of grade three, i.t.a. pupils of <u>all</u> ability levels were able to spell as well as the T.O. pupils.

In spite of the different research designs, different pupils, different teachers, and different reading programs involved in the various i.t.a. studies, the pattern of research results, including those in this study, remains consistent. When tested in i.t.a. during grade one, the i.t.a. pupils demonstrated significant superiority over the T.O. pupils in skills which require sound-letter association, such as Word Reading and Word Study Skills. In the post-transition period in grades two and three, few significant differences were found between i.t.a. pupils and T.O. pupils on spelling and other reading subtests. The concern of some educators that i.t.a. might have long-lasting and detrimental effects on pupils' T.O. spelling and reading has not been substantiated by research.

This study further suggests that the type of program used in the post-transition period is a key factor to success in T.O. reading, particularly for the low ability pupils. In most cases, the scores of low ability i.t.a. pupils who used T.O. programs emphasizing phonics averaged at or above grade placement. This suggests that a phonics-emphasized program probably should be used in order to capitalize on the early advantages achieved through the use of i.t.a. Further study of this topic is needed.

The study also suggests that the introduction of i.t.a. in kindergarten is not only feasible, but necessary if pupils are to show benefits beyond grade one. The group of pupils who started i.t.a. instruction in kindergarten maintained their advantage through the end of grade three. Since no T.O. classes introducing reading instruction in kindergarten were available to this study, it is not possible to compare the merits of starting T.O. or i.t.a. in kindergarten.

The use of i.t.a. has been questioned by some educators because of the lack of statistical differences in reading achievement between i.t.a. and T.O. pupils beyond grade one in most studies. If it is assumed that we want children to have the best possible success at every level of school learning, i.t.a. certainly has its merits in enhancing the beginning reader's word attacking skills. Furthermore, since i.t.a. spelling is regular, the child is likely to be able to read the i.t.a. books with relatively little help from the teacher once he learns to decode. This means that the i.t.a. books need not be as limited in vocabulary, scope, and content as the T.O. books for the beginning readers. Thus, the child could be helped to develop an enthusiasm for reading and for learning with a wide range of interesting and educationally valuable materials. Attention, however, should be given to utilizing the i.t.a. pupils' superior word study skills to develop vocabulary and comprehension.

PART 2

TECHNICAL REPORT

The importance of success in beginning reading has long been recognized by those interested in education. The best way to achieve this success remains controversial. The initial teaching alphabet, i.t.a., was designed with the purpose of simplifying the learning task for the beginners. Forty-four characters are used to increase the consistency between written symbols and sounds in the English language. Further simplification is realized through elimination of the capitals.

The effects of i.t.a. on reading achievement have been the subject of investigation in many studies. Indeed, few educational innovations can claim as much research. Six major studies have been reported since its introduction to American education in 1963 (Tanyzer & Alpert, 1966; Chasnoff, 1967, 1968; Fry, 1966, 1967a, 1969; Hayes, 1966; Hayes & Wuest, 1967, 1969; Hahn, 1966, 1967; Mazurkiewicz, 1966, 1967). The pupils instructed in i.t.a. were usually compared with those taught with various reading programs in T.O., the traditional orthography. In spite of the use of different pupils, different teachers, different reading programs, and different research designs which have often been criticized (Asher, 1968; Block, 1966; Fry, 1967b; Gillooly, 1967, 1968), the trend of the findings is highly consistent and predictable. When tested in their own instructional medium in grade one, the i.t.a. pupils performed as well as the T.O. pupils in vocabulary and reading comprehension, but excelled significantly their T.O. counterparts in tests which require symbol-sound association, such as word reading and word study skills. When tested in T.O. at or beyond the end of grade one, seldom were any significant differences found between the T.O. and i.t.a. pupils on the standardized tests.

The study reported here was an attempt to provide answers to some questions not yet explored in earlier studies. Two major purposes were embodied in this study: investigation of (a) longitudinal effects of i.t.a. on reading achievement of pupils of various ability levels; (b) feasibility of starting i.t.a. instruction in kindergarten.

Procedure

Subjects

Seven hundred fifteen pupils from nine Council school districts participated in the study. The same pupils were followed through the end of grade three. Of this sample, 281 pupils started T.O. in grade one, 213 pupils started i.t.a. in grade one, and 221 pupils started i.t.a. in kindergarten. The three groups will be designated the T.O. group, the i.t.a.-grade one (or i.t.a.-1) group, and the i.t.a.-kindergarten (or i.t.a.-k)

group respectively. In all the school districts, i.t.a. was implemented two years before this study began.

Instructional Materials

All i.t.a. pupils used the Early To Read series, supplemented by the Downing i.t.a. program in some cases. They generally transferred to T.O. during the second semester in grade two. The T.O. readers used in the post-i.t.a. period were mainly the Ginn; Scott, Foresman; and Lippincott series. Most i.t.a. pupils were kept intact in the same classrooms in the first two grades. The T.O. pupils used a great variety of basal reading series including those published by: Scott, Foresman; Ginn; Harper-Row; Row-Peterson; Houghton Mifflin; American Book Company; Lippincott; etc.

Variables

The Lorge-Thorndike Intelligence Tests, Level 1, Form B, were given in the third month of the first grade. The reading achievement of the pupils was measured by the subtests in the Stanford Achievement Test: Word Reading, Paragraph Meaning, Vocabulary, Spelling, and Word Study Skills for grade one; Word Meaning, Paragraph Meaning, Spelling, Word Study Skills, and Language for grades two and three. In the fifth month of grade one, the i.t.a. pupils took the tests in i.t.a., while the T.O. pupils took the corresponding tests in T.O. At the end of grades one, two, and three, all pupils were tested in T.O.

At the end of grade three, teachers were asked to rate each pupil on four five-point scales: overall school adjustment, ability to apply oneself to learning, attitude toward reading, and emotional adjustment.

Analysis

The pupils were classified into three ability levels on the basis of the Lorge-Thorndike IQ's. Each ability level included approximately one-third of the 715 pupils. The mean IQ was 95 for the low level; 108 for the middle level; and 121 for the high level. Within each ability level, the pupils were further subdivided by treatment. The number of pupils, mean IQ, and IQ range for each subgroup are given in Table 1. These data show that pupils of different treatments within each ability level were comparable with respect to IQ.

Table 1

Number of pupils, mean IQ, and IQ range

within each treatment group by ability level

Ability level	Treatment	Number of pupils	Mean IQ by treatment	IQ range by ability level	Mean IQ by ability level
High	i.t.a.-1	40	120	115-139	121
	T.O.	99	122		
	i.t.a.-k	90	122		
Middle	i.t.a.-1	75	108	104-114	108
	T.O.	90	108		
	i.t.a.-k	77	108		
Low	i.t.a.-1	98	94	64-103	95
	T.O.	92	95		
	i.t.a.-k	54	94		

The Hierarchal Analysis[1] and the Newman-Keuls Test

In order to account for the unique effects associated with each classroom, the hierarchal analysis was used to test the differences among the treatment groups for each dependent variable[2] at each ability level. A significant F value may thus be explained as due to the treatment rather than to the combination of treatment and classroom effects. This analysis involved the nesting of classrooms within each treatment. To accomplish the nesting at grades two and three, the class was arbitrarily assigned to the treatment whose pupils were in the majority; pupils belonging to other treatments were deleted. The one-pupil classes which resulted from ability grouping and nesting were also dropped, since computation of the within class variance was impossible. As a result, the analysis involved different numbers of pupils from grade to grade. The actual numbers of pupils included in the analysis

---

[1] Dr. Maurice Tatsuoka of the University of Illinois made the adjustments in the general model for unequal numbers of students within each classroom and unequal numbers of teachers nested under each treatment (see Appendix A).

[2] The grade one T.O. Spelling test given at 1.9 grade placement was not included in the comparison. Since most i.t.a. pupils were not able to spell in T.O. at the end of grade one, it is felt that the comparison was meaningless.

and the IQ means are shown in Table 2.[3] The deletions (from the original 715 pupils) caused little change in the IQ's of the groups from year to year.

When the F value of the hierarchal analysis was significant, the Newman-Keuls test was used to compare each pair of the treatment means.

Table 2

Size and mean IQ's of samples on which analysis was done

Ability level	Treatment	Grade 1		Grade 2		Grade 3	
		N	Mean IQ	N	Mean IQ	N	Mean IQ
High	i.t.a.-1	35	120	32	120	27	119
	T.O.	97	122	89	122	82	122
	i.t.a.-k	89	122	86	122	88	122
Middle	i.t.a.-1	72	108	65	108	56	108
	T.O.	88	108	79	108	72	108
	i.t.a.-k	76	108	77	108	76	108
Low	i.t.a.-1	89	94	79	94	68	94
	T.O.	88	95	85	96	67	96
	i.t.a.-k	50	94	50	94	49	94

Analysis of the Use of the T.O. Programs in the Post-i.t.a. Period

This analysis is of a descriptive nature. In order to detect the possible pattern that might exist among classes that used different types of T.O. programs in the post-i.t.a. period, the T.O. program used by each class was identified and classified under the "meaning emphasized" and "phonics emphasized" categories according to Chall's guidelines (1967). The

---

[3] It should be noted that the number of pupils for grade one in this table is not the same as that in Table 1. This is because pupils who did not have complete data in grades two and three were not included in the analysis.

"meaning emphasized" category included the conventional basal readers such as the Ginn and Scott, Foresman series. The "phonics emphasized" program(s) referred to the Lippincott series, either used exclusively or supplemented by other T.O. series. Programs whose classification was not clear were classified under the category "others." The mean scores of the classes under each category on the reading tests were plotted by ability level for grades two and three.

Results

The results of the hierarchal analysis are presented in Tables 3-5 which correspond to the high, middle, and low ability samples respectively. These tables give the means of each treatment group on each dependent variable, the F values of the tests of the differences among treatments, and the F values of the tests of differences among classrooms. Results of the Newman-Keuls tests are given in Table 6. Mean scores of classes using various T.O. programs in grades two and three are plotted in Figures 1-6.

The major findings of the study can be summarized as follows:

1. In the fifth month of grade one, when the treatment groups were tested in their own instructional medium, the i.t.a.-kindergarten group demonstrated significant superiority to the T.O. group in Word Reading, Paragraph Meaning, Spelling, and Word Study Skills at all three ability levels. The i.t.a.-grade one group scored significantly higher than the T.O. group on the following tests: Word Reading, Paragraph Meaning, Spelling, and Word Study Skills at the high ability level; Word Reading, Spelling, and Word Study Skills at the middle ability level; Word Reading and Word Study Skills at the low ability level.

2. At the end of grade one, when the subtests on Word Reading, Paragraph Meaning, Vocabulary, and Word Study Skills were given in T.O., no significant differences were found between the i.t.a. groups and the T.O. group at all ability levels. Most of the i.t.a. pupils had not made the formal transition to T.O. at that time.

3. Beyond grade one, no significant differences were found between the i.t.a.-1 group and the T.O. group on Word Meaning, Paragraph Meaning, Word Study Skills, and Language at all ability levels. On the Spelling subtest, middle ability T.O. pupils scored significantly higher than their i.t.a.-1 counterparts at the end of grade two. No significant spelling differences were found between the i.t.a.-1 and T.O. high and low ability groups at the end of grade two or between the i.t.a.-1 and T.O. pupils of any ability group at the end of grade three.

Table 3

Mean scores, teacher ratings, and the results

of hierarchal analyses of variance

(High ability pupils)

Dependent variable	Treatment means			Treatment effect		Classroom effect	
	i.t.a.-1	T.O.	i.t.a.-k	df	F	df	F
SAT, Word Reading (1.5)[a]	2.46[b]	1.77	2.82	2,40	32.45**	36,182	3.11**
SAT, Paragraph Meaning(1.5)	2.04	1.74	2.29	2,40	6.87**	36,182	3.21**
SAT, Vocabulary (1.5)	2.58	2.67	2.90	2,42	1.17	36,182	2.06**
SAT, Spelling (1.5)	2.39	1.81	2.56	2,41	28.07**	36,182	2.41**
SAT, Word Study Skills (1.5)	3.90	2.20	4.00	2,43	36.91**	36,182	1.96**
SAT, Word Reading (1.9)	2.44	2.31	2.53	2,41	1.36	36,182	2.57**
SAT, Paragraph Meaning(1.9)	2.55	2.25	2.45	2,41	1.35	36,182	2.52**
SAT, Vocabulary (1.9)	3.06	2.96	3.15	2,42	0.53	36,182	2.10**
SAT, Word Study Skills (1.9)	3.40	2.78	3.19	2,43	2.72	36,182	1.84**
SAT, Word Meaning (2.8)	3.67	3.65	3.97	2,43	1.27	37,167	2.41**
SAT, Paragraph Meaning(2.8)	3.61	3.79	4.25	2,44	3.15	37,167	2.09**
SAT, Spelling (2.8)	3.36	3.53	3.67	2,44	0.56	37,167	2.19**
SAT, Word Study Skills (2.8)	4.69	4.37	5.36	2,46	4.92*	37,167	1.78**
SAT, Language (2.8)	4.03	3.80	4.03	2,42	0.40	37,167	3.09**
SAT, Word Meaning (3.9)	4.63	4.86	5.24	2,51	1.78	42,152	2.37**
SAT, Paragraph Meaning(3.9)	4.44	4.79	5.12	2,52	2.07	42,152	2.15**
SAT, Spelling (3.9)	4.20	4.52	4.54	2,54	0.57	42,152	1.97**
SAT, Word Study Skills (3.9)	5.14	5.13	6.17	2,55	6.84**	42,152	1.52**
SAT, Language (3.9)	5.04	4.99	5.09	2,57	0.06	42,152	1.53*
T.R. School Adjustment	4.0	4.2	4.1	2,57	0.28	42,152	1.53*
T.R. Application to Learning	4.1	3.9	4.1	2,56	0.44	42,152	1.72*
T.R. Attitude Toward Reading	4.1	4.1	4.1	2,53	0.01	42,152	2.01**
T.R. Emotional Adjustment	4.0	4.1	4.0	2,60	0.21	42,152	1.35

[a] The figures in parentheses refer to the pupils' grade placement at the time of testing.

[b] All mean scores on the Stanford reading subtests are grade equivalent scores.

* $p<.05$.

** $p<.01$.

Table 4

Mean scores, teacher ratings, and the results

of hierarchal analyses of variance

(Middle ability pupils)

Dependent variable	Treatment means			Treatment effect		Classroom effect	
	i.t.a.-1	T.O.	i.t.a.-k	df	F	df	F
SAT, Word Reading (1.5)[a]	2.17[b]	1.68	2.62	2,46	25.27**	41,192	3.14**
SAT, Paragraph Meaning (1.5)	1.67	1.58	2.06	2,47	11.35**	41,192	2.56**
SAT, Vocabulary (1.5)	2.31	2.16	2.35	2,49	0.84	41,192	1.97**
SAT, Spelling (1.5)	2.05	1.74	2.43	2,46	12.82**	41,192	3.25**
SAT, Word Study Skills (1.5)	2.94	2.04	3.59	2,49	18.83**	41,192	2.02**
SAT, Word Reading (1.9)	2.05	2.18	2.32	2,47	1.96	41,192	2.58**
SAT, Paragraph Meaning (1.9)	1.93	2.09	2.30	2,47	2.60	41,192	2.67**
SAT, Vocabulary (1.9)	2.66	2.53	2.64	2,49	0.40	41,192	1.99**
SAT, Word Study Skills (1.9)	2.51	2.65	2.77	2,46	0.42	41,192	2.95**
SAT, Word Meaning (2.8)	3.23	3.41	3.56	2,51	0.94	44,174	3.12**
SAT, Paragraph Meaning (2.8)	3.17	3.37	3.71	2,53	2.62	44,174	2.43**
SAT, Spelling (2.8)	2.98	3.44	3.57	2,57	4.52*	44,174	1.74**
SAT, Word Study Skills (2.8)	3.93	4.41	4.67	2,56	1.68	44,174	1.87**
SAT, Language (2.8)	3.18	3.50	3.62	2,51	1.34	44,174	3.08**
SAT, Word Meaning (3.9)	4.38	4.56	4.65	2,57	0.47	44,157	1.95**
SAT, Paragraph Meaning (3.9)	4.18	4.50	4.56	2,62	1.61	44,157	1.43
SAT, Spelling (3.9)	4.22	4.43	4.53	2,63	1.15	44,157	1.32
SAT, Word Study Skills (3.9)	4.88	5.03	5.73	2,60	3.36*	44,157	1.56*
SAT, Language (3.9)	4.40	4.68	4.74	2,58	0.75	44,157	1.77**
T.R. School Adjustment	3.8	3.9	3.9	2,55	0.19	44,157	2.19**
T.R. Application to Learning	3.6	3.8	3.9	2,64	0.90	44,157	1.28
T.R. Attitude Toward Reading	3.7	3.9	4.0	2,60	0.99	44,157	1.54*
T.R. Emotional Adjustment	3.6	3.9	3.9	2,61	1.87	44,157	1.51*

[a] The figures in parentheses refer to the pupils' grade placement at the time of testing.

[b] All mean scores on the Stanford reading subtests are grade equivalent scores.

* $p < .05$.

** $p < .01$.

Table 5

Mean scores, teacher ratings, and the results of hierarchal analyses of variance

(Low ability pupils)

Dependent variable	Treatment means			Treatment effect		Classroom effect	
	i.t.a.-1	T.O.	i.t.a.-k	df	F	df	F
SAT, Word Reading (1.5)[a]	1.94[b]	1.46	2.39	2,43	19.28**	39,185	3.63**
SAT, Paragraph Meaning (1.5)	1.59	1.49	1.85	2,44	6.07**	39,185	3.14**
SAT, Vocabulary (1.5)	1.92	1.85	2.11	2,48	1.83	39,185	1.72*
SAT, Spelling (1.5)	1.74	1.53	2.19	2,41	6.64**	39,185	5.07**
SAT, Word Study Skills (1.5)	2.65	1.75	3.16	2,42	9.30**	39,185	4.15**
SAT, Word Reading (1.9)	1.94	1.89	2.16	2,45	2.57	39,185	2.55**
SAT, Paragraph Meaning (1.9)	1.79	1.82	1.99	2,44	1.17	39,185	2.94**
SAT, Vocabulary (1.9)	2.12	2.19	2.27	2,47	0.61	39,185	1.98**
SAT, Word Study Skills (1.9)	2.24	2.06	2.52	2,46	2.63	39,185	2.07**
SAT, Word Meaning (2.8)	3.05	2.94	3.18	2,55	0.69	47,164	3.01**
SAT, Paragraph Meaning (2.8)	2.98	2.88	3.17	2,57	0.90	47,164	2.56**
SAT, Spelling (2.8)	2.95	3.03	3.33	2,56	1.15	47,164	2.86**
SAT, Word Study Skills (2.8)	3.73	3.43	3.81	2,56	0.51	47,164	2.72**
SAT, Language (2.8)	3.03	3.06	3.22	2,57	0.47	47,164	2.49**
SAT, Word Meaning (3.9)	3.86	3.87	4.02	2,58	0.21	44,137	1.97**
SAT, Paragraph Meaning (3.9)	3.73	3.73	3.90	2,54	0.24	44,137	2.71**
SAT, Spelling (3.9)	3.86	3.97	4.17	2,53	0.48	44,137	2.98**
SAT, Word Study Skills (3.9)	4.54	4.24	4.86	2,60	1.00	44,137	1.81**
SAT, Language (3.9)	4.02	3.91	3.98	2,56	0.06	44,137	2.39**
T.R. School Adjustment	3.6	3.7	3.5	2,61	0.33	44,137	1.71*
T.R. Application to Learning	3.4	3.6	3.3	2,74	1.30	44,137	1.00
T.R. Attitude Toward Reading	3.5	3.7	3.6	2,59	0.25	44,137	1.88**
T.R. Emotional Adjustment	3.4	3.8	3.6	2,64	1.81	44,137	1.48

[a] The figures in parentheses refer to the pupils' grade placement at the time of testing.

[b] All mean scores on the Stanford reading subtests are grade equivalent scores.

* $p<.05$.

** $p<.01$.

Table 6

Results of the Newman-Keuls tests for dependent variables whose treatment effects were significant in the hierarchal analysis of variance

Ability level	Dependent variable	i.t.a.-1 vs. T.O.			i.t.a.-k vs. i.t.a.-1			i.t.a.-k vs. T.O.		
		i.t.a.-1 Mean	T.O. Mean	Critical Value[a]	i.t.a.-k Mean	i.t.a.-1 Mean	Critical Value	i.t.a.-k Mean	T.O. Mean	Critical Value
Low	Word Reading (1.5)[b]	1.94[c]	1.46	.297*	2.39	1.94	.297*	2.39	1.46	.359*
	Paragraph Meaning (1.5)	1.59	1.49	.202	1.85	1.59	.202*	1.85	1.49	.244*
	Spelling (1.5)	1.74	1.53	.354	2.19	1.74	.354*	2.19	1.53	.428*
	Word Study Skills (1.5)	2.65	1.75	.678*	3.16	2.65	.678	3.16	1.75	.819*
Middle	Word Reading (1.5)	2.17	1.68	.276*	2.62	2.17	.276*	2.62	1.68	.333*
	Paragraph Meaning (1.5)	1.67	1.58	.219	2.06	1.67	.219*	2.06	1.58	.264*
	Spelling (1.5)	2.05	1.74	.276*	2.43	2.05	.276*	2.43	1.74	.333*
	Word Study Skills (1.5)	2.94	2.04	.532*	3.59	2.94	.532*	3.59	2.04	.642*
	Spelling (2.8)	2.98	3.44	.407*	3.57	2.98	.491*	3.57	3.44	.407
	Word Study Skills (3.9)	4.88	5.03	.748	5.73	4.88	.847*	5.73	5.03	.748
High	Word Reading (1.5)	2.46	1.77	.332*	2.82	2.46	.332	2.82	1.77	.401*
	Paragraph Meaning (1.5)	2.04	1.74	.379	2.29	2.04	.379	2.29	1.74	.457*
	Spelling (1.5)	2.39	1.81	.261*	2.56	2.39	.261	2.56	1.81	.315*
	Word Study Skills (1.5)	3.90	2.20	.567*	4.00	3.90	.567	4.00	2.20	.685*
	Word Study Skills (2.8)	4.69	4.37	.838	5.36	4.69	.838	5.36	4.37	.988*
	Word Study Skills (3.9)	5.14	5.13	.720	6.17	5.14	.720*	6.14	5.13	.875*

[a] The critical value of the difference between two means is $\sqrt{\text{MS error}/\overline{n}}\ q_{1-\alpha}$, (r, df) where $q_{1-\alpha}$, (r, df) is the studentized range statistic.

[b] The figures in parentheses refer to the pupils' grade placement at the time of testing.

[c] All mean scores on the Stanford reading subtests are grade equivalent scores.

* $p < .05$.

FIGURE 1
MEAN SCORES OF CLASSES USING THE VARIOUS T.O. PROGRAMS
BY THE TREATMENT GROUPS IN GRADE TWO
(HIGH ABILITY GROUP)

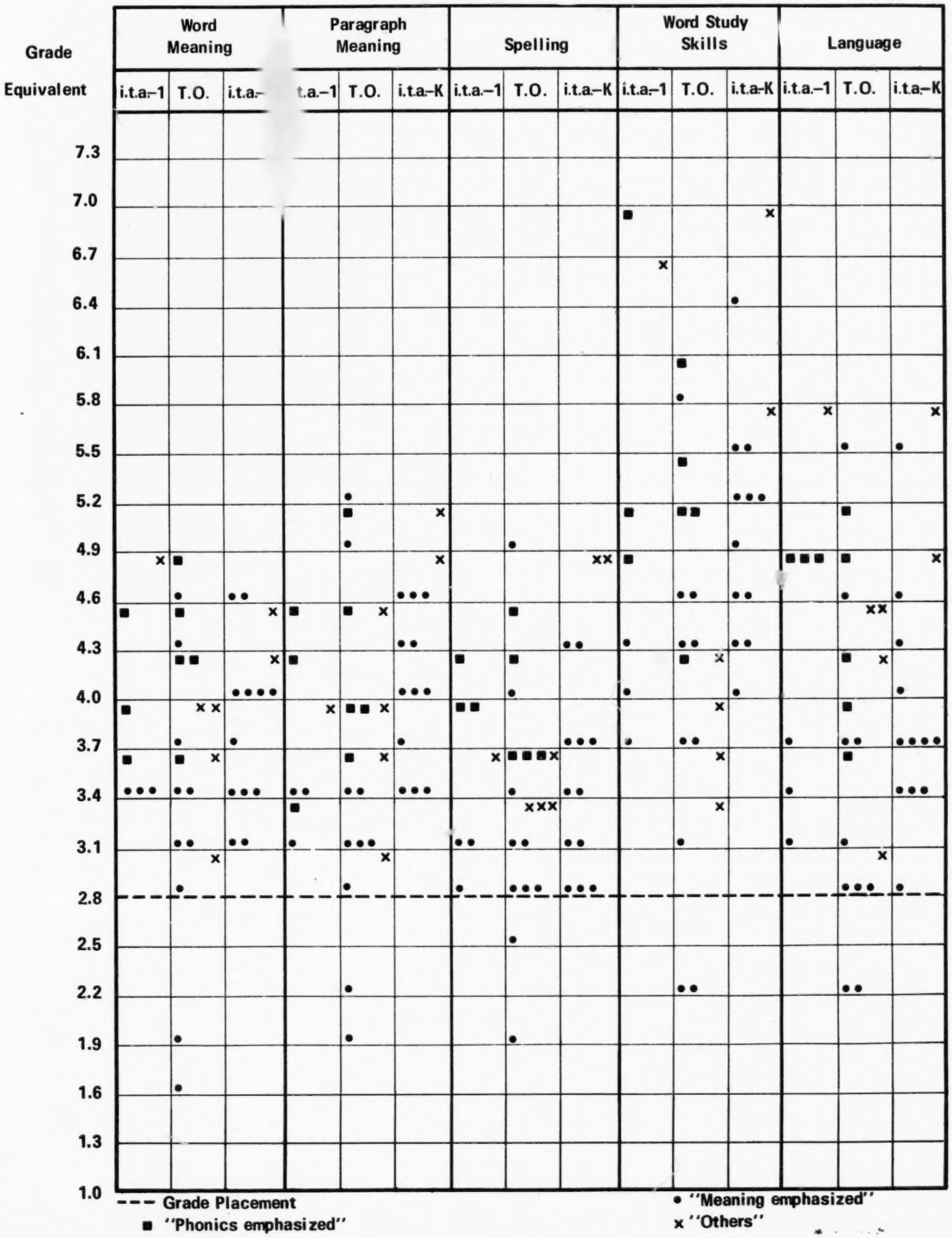


FIGURE 2
MEAN SCORES OF CLASSES USING THE VARIOUS T.O. PROGRAMS
BY THE TREATMENT GROUPS IN GRADE TWO
(MIDDLE ABILITY GROUP)

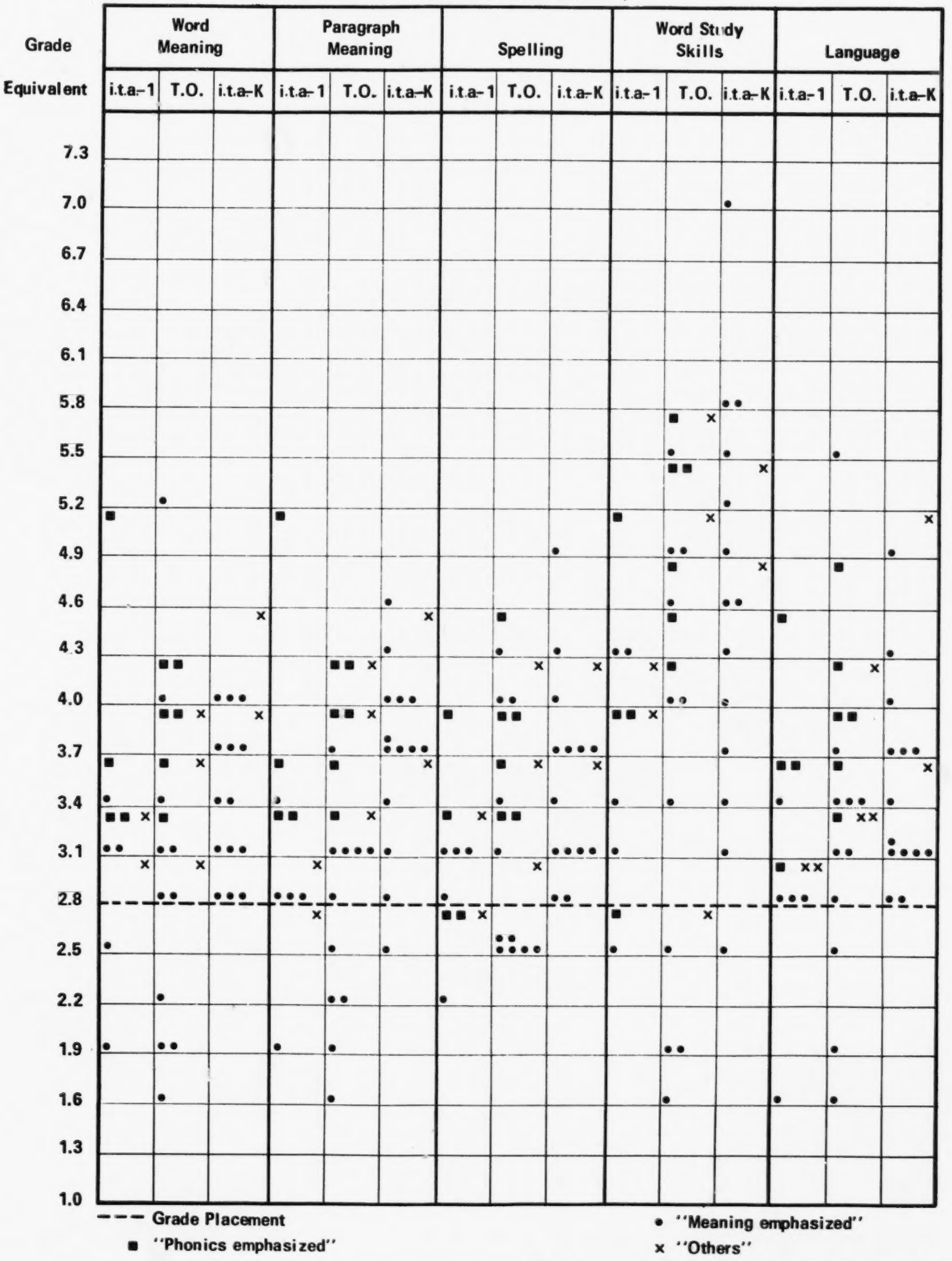


FIGURE 3
MEAN SCORES OF CLASSES USING THE VARIOUS T.O. PROGRAMS
BY THE TREATMENT GROUPS IN GRADE TWO
(LOW ABILITY GROUP)

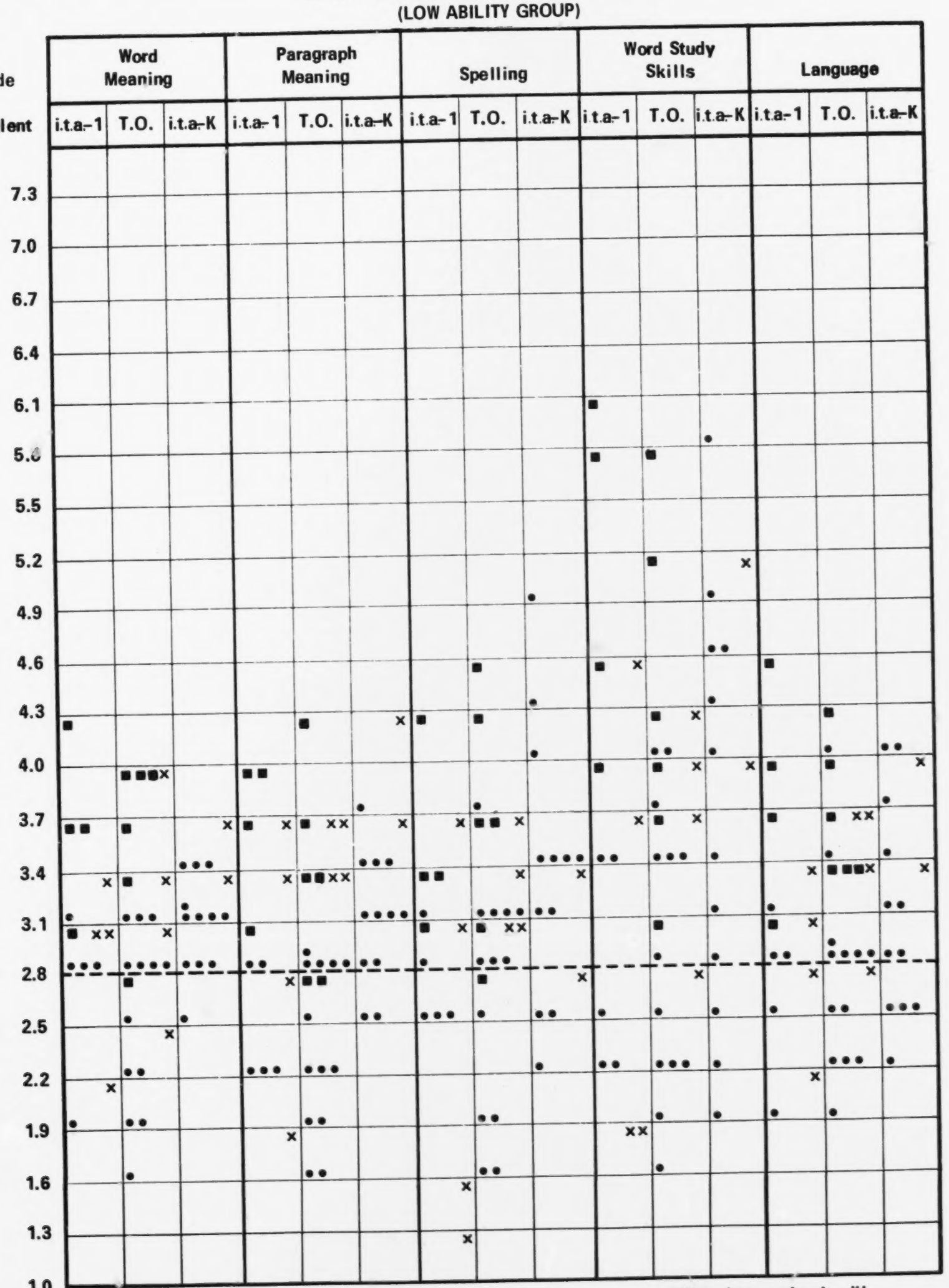


FIGURE 4
MEAN SCORES OF CLASSES USING THE VARIOUS T.O. PROGRAMS
BY THE TREATMENT GROUPS IN GRADE THREE
(HIGH ABILITY GROUP)

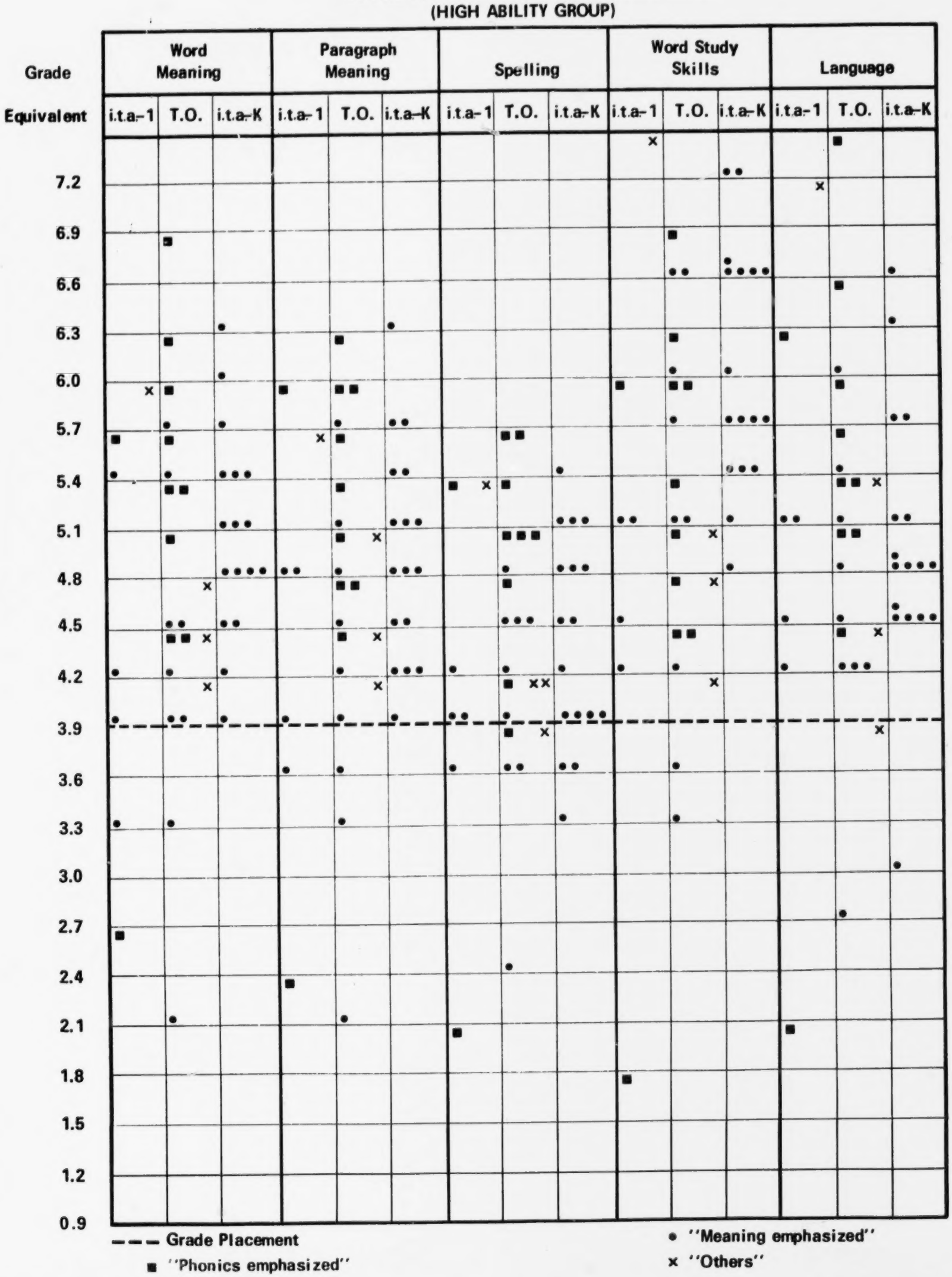


FIGURE 5
MEAN SCORES OF CLASSES USING THE VARIOUS T.O. PROGRAMS
BY THE TREATMENT GROUPS IN GRADE THREE
(MIDDLE ABILITY GROUP)

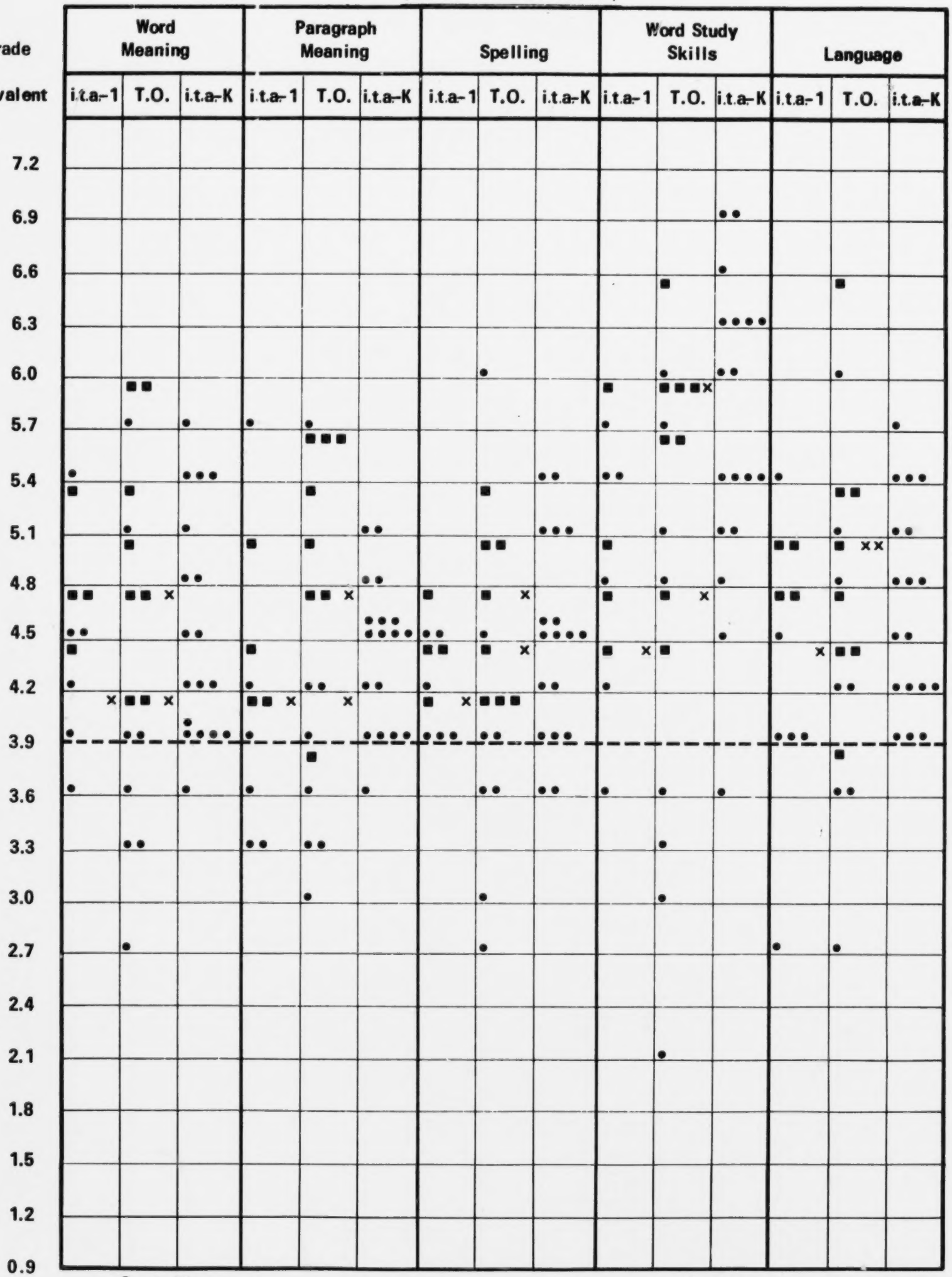
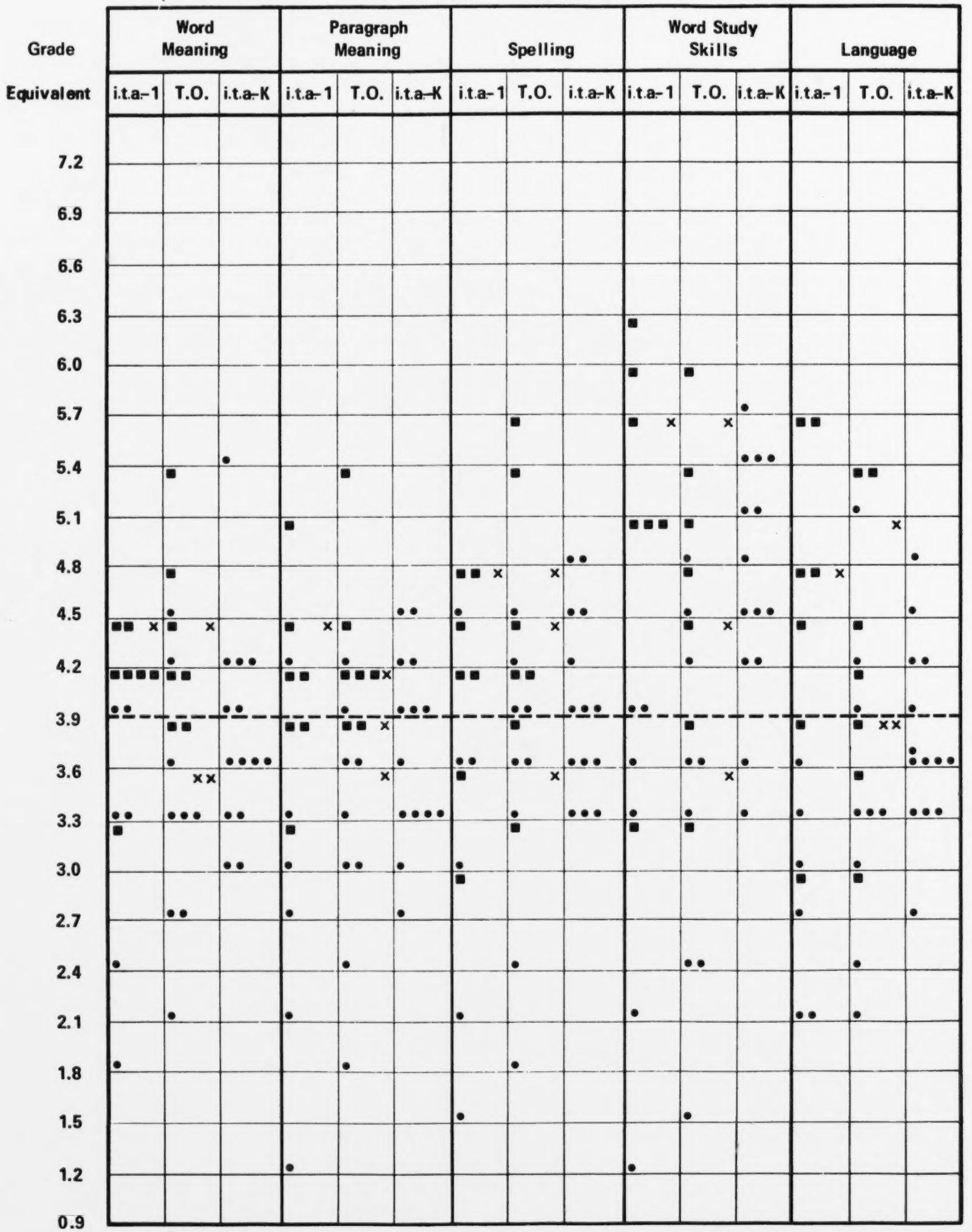


FIGURE 6
MEAN SCORES OF CLASSES USING THE VARIOUS T.O. PROGRAMS
BY THE TREATMENT GROUPS IN GRADE THREE
(LOW ABILITY GROUP)



4. Beyond grade one, significant differences in Word Study Skills were due to the superiority of the middle and high ability i.t.a.-k pupils over either one or both of the corresponding i.t.a.-1 and T.O. pupils. The i.t.a.-kindergarten group also outperformed the other two groups on almost all the subtests throughout the three years at all ability levels even though few statistical significances were obtained.

5. Those i.t.a. pupils who used a reading series that emphasized the "phonics approach" tended to be superior to their counterparts who used readers that emphasized the "meaning approach." Most of the mean scores of the low ability level in the former group were at or above the grade placement for all subtests at the end of grades two and three.

6. The analysis of third grade teacher ratings of pupils' school adjustment, ability to apply self to learning, attitude toward reading, and general emotional adjustment revealed no significant differences among the i.t.a. and T.O. groups at all ability levels. However, the data indicate that teachers' ratings on each of these characteristics vary according to the ability of the child. High ability children were rated as having better school adjustment, greater ability to apply themselves, better attitudes toward reading, and better emotional adjustment than middle and low ability children regardless of reading program.

7. Classroom effects were significant for virtually all reading subtests and teacher ratings at all ability levels.

Discussion

The results of this study were consistent with those of earlier studies even when the classroom effects and ability levels were taken into consideration. When tested in i.t.a. during grade one, the i.t.a. pupils demonstrated significant superiority over the T.O. pupils in skills which require sound-letter association, such as Word Reading and Word Study Skills. When tested in T.O. at the end of grades two and three, few significant differences were found between i.t.a. pupils and T.O. pupils on spelling and other reading subtests. The concern of some educators that i.t.a. might have adverse effects on pupils' T.O. spelling and reading has not been substantiated by research.

The possibility that some kind of automatic transition from i.t.a. to T.O. might occur is suggested by the fact that no significant differences were found between the i.t.a. groups and T.O. group at the end of grade one before most i.t.a. pupils had made the formal transition. This and other questions about how children make the transition from i.t.a. to T.O. should be more fully investigated.

This study suggests that the type of program used in the post-transition period is a key factor to success in T.O. reading, particularly for the low ability pupils. The data reflect a trend favoring a phonics-oriented program. This kind of program probably helps to capitalize on the early advantages achieved through the use of i.t.a. Further study on this topic should be conducted.

This study shows that starting i.t.a. in kindergarten is not only possible, but seems to have long-term beneficial effects on pupils' achievement. The group of pupils who started i.t.a. instruction in kindergarten maintained their advantage through the end of grade three. Since no T.O. classes beginning reading instruction in kindergarten were available for this study, it is not possible to compare the merits of starting T.O. or i.t.a. in kindergarten. Further study is needed to provide the answer.

The fact that there were no significant differences between the T.O. and i.t.a. groups on the teachers' ratings might suggest the following: (a) The instruments were too crude. (b) Teachers' interpretations of the rating scales were different. (c) The difference between the T.O. and i.t.a. groups might occur in grades one and two but disappear in grade three. (d) The limited categories in the five-point ratings might make it impossible to discriminate among the treatment groups within each ability level. Greater emphasis should be given to the development of scales which will measure affective variables.

The use of i.t.a. has been questioned by some educators because of the lack of statistical differences in reading achievement between i.t.a. and T.O. pupils beyond grade one in most studies. If it is assumed that we want children to have the best possible success at every level of school learning, i.t.a. certainly has its merits in enhancing the beginning reader's word attacking skills. Furthermore, since i.t.a. spelling is regular, the child is likely to be able to read the i.t.a. books with relatively little help from the teacher once he learns to decode. This means that the i.t.a. books need not be as limited in vocabulary, scope, and content as the T.O. books for the beginning readers. Thus, the child could be helped to develop an enthusiasm for reading and for learning with a wide range of interesting and educationally valuable materials. Attention, however, should be given to utilizing the i.t.a. pupils' superior word study skills to develop vocabulary and comprehension.

APPENDIX A

HIERARCHAL ANALYSIS OF VARIANCE WITH UNEQUAL SAMPLE SIZES

Maurice M. Tatsuoka

As described in the main body of this report, a two-factor hierarchal design was applied to each of nine samples of pupils, defined in terms of two descriptor variables: Grade Level (first, second, and third) and Ability Level (high, average, and low).

The two factors (or independent variables) were: [A] Type of Initial Reading Program (i.t.a-K, i.t.a.-1, and T.O.) and [B] Teacher. The design is said to be hierarchal or "nested," because different levels of Factor B (i.e. different teachers) occur within the three levels of Factor A. A schematic representation of the design is, therefore, as follows:

A_1 (i.t.a.-K)				A_2 (i.t.a.-1)			A_3 (T.O.)		
B_1	B_2	B_3 ...	B_{16}	B_{17}	B_{18} ...	B_{24}	B_{25}	B_{26} ...	B_{39}
x	x	x	x	x	x	x	x	x	x
x	x	x	x	x	x	x	.	x	x
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	x	.	.
x	.	x	x	.	x	.		.	.
	.			.		x		x	.
	x			x					x

This means that there were 39 teachers in all (B_1, B_2, ...,B_{39}), 16 of whom taught pupils with the A_1 initial reading program, 8 of whom taught A_2 pupils, and 15 taught A_3 pupils. (The actual numbers are those for the grade one high ability sample. The other eight samples will have different ranges of the

subscript for B under each A level.) The x's represent the observations, or dependent-variable values (scores) for pupils taught by each teacher.

The idea of hierarchy or nestedness may perhaps be further clarified by comparing the above design (the one actually used in this study) with a design in which nesting does not occur. The latter would apply if it had been the case that the same teachers (say 13 in number) taught pupils of all three initial reading program types (i.e. levels of Factor A), thus:

A_1				A_2				A_3			
B_1	B_2	...	B_{13}	B_1	B_2	...	B_{13}	B_1	B_2	...	B_{13}
x	x		x	x	x		x	x	x		x
x	x		x	.	x		.	x	x		.
.	.		.	.	.		.	.	.		.
.	.		.	.	.		.	.	.		x
.	x		.	x	.		.	.	.		
x			x		.		x	x	.		
					x				x		

Note that, although we again have 39 groups of pupils, we now do not have 39 teachers but only 13; the same 13 teachers occur under all three levels of Factor A. In this case, Factor B is not nested within Factor A, but is said to be crossed with Factor A.

Coming back to the hierarchal design that was actually used in this study, recall that there were several dependent variables for each grade level. More specifically, Tables 3-5[1] show that there were ten dependent variables

[1] See pp. 9-11 in the Technical Report.

for grade one, five for grade two, and nine for grade three. Since there were three ability levels in each grade, this means that 30 + 15 + 27 = 72 separate two-factor hierarchal-design analyses were carried out in all. The outline and formulas presented below apply to each one of these 72 separate analyses.

To simplify the notation somewhat, we now switch from the consecutive numbering (B_1, B_2, ..., B_{39}) of levels of Factor B (teachers) to a system which enumerates the teachers separately for each level of Factor A. That is, we henceforth designate the B levels by a double-subscript notation, thus, for example:

$B_{1(1)}, B_{2(1)}, \ldots, B_{16(1)}$ for teachers nested in A_1;

$B_{1(2)}, B_{2(2)}, \ldots, B_{8(2)}$ for teachers nested in A_2;

$B_{1(3)}, B_{2(3)}, \ldots, B_{15(3)}$ for teachers nested in A_3.

Thus, the teacher previously denoted by B_{12} is now represented by $B_{12(1)}$; teacher B_{20} is now $B_{4(2)}$; B_{32} becomes $B_{8(3)}$; and so on. In general, $B_{j(i)}$ denotes the j-th teacher nested in A_i, the i-th level of Factor A (where i = 1, 2, 3). (This may seem like a complication rather than simplification of notation, but it actually simplifies the notation in the subsequent formulas!) More generally, we denote by β_i the number of teachers nested in A_i (i = 1, 2, 3).

With the foregoing notational conventions, we are now ready to present the formulas for the various sums-of-squares (SS's), their degrees of freedom (df's), and the resulting mean-squares (MS's) that enter into the significance tests. We denote by X, the dependent variable used in any one of the 72 analyses carried out. Triple subscripts are needed to specify individual scores on X, as follows:

X_{ijk} = score of k-th pupil in the class taught by j-th teacher nested in A_i.

The subscripts here have the following ranges:

$i = 1, 2, 3$ (there being three programs, A_1, A_2, A_3),

$j = 1, 2, \ldots, \beta_i$ (= the number of teachers nested in A_i),

$k = 1, 2, \ldots, n_{ij}$ (= the number of pupils in teacher $B_{(j)i}$'s class).

The several kinds of means are defined and denoted as follows:

$$\bar{X}_{ij.} = \left(\sum_{k=1}^{n_{ij}} X_{ijk} \right) \Big/ n_{ij} \quad \text{(the mean for teacher } B_{(j)i}\text{'s class)},$$

$$\bar{X}_{i..} = \left(\sum_{j=1}^{\beta_i} \sum_{k=1}^{n_{ij}} X_{ijk} \right) \Big/ N_i \quad \text{(the mean for program } A_i\text{)},$$

$$\text{where } N_i = \sum_{j=1}^{\beta_i} n_{ij} \quad \text{(the number of pupils in } A_i\text{)}$$

$$\bar{X} = \left(\sum_{i=1}^{3} \sum_{j=1}^{\beta_i} \sum_{k=1}^{n_{ij}} X_{ijk} \right) \Big/ N \quad \text{(the grand mean),}$$

$$\text{where } N = \sum_{i=1}^{3} N_i \quad \text{(total size of any grade-ability sample).}$$

Finally, the SS's and their df's are given by the formulas:

$$\text{SS within B} = \sum_{i=1}^{3} \sum_{j=1}^{\beta_i} \sum_{k=1}^{n_{ij}} \left(X_{ijk} - \bar{X}_{ij.} \right)^2,$$

$$\text{df} = \sum_{i,j} \left(n_{ij} - 1 \right) = N - \left(\beta_1 + \beta_2 + \beta_3 \right);$$

$$SS_{B(A)} = \sum_{i=1}^{3} \sum_{j=1}^{\beta_i} n_{ij} \left(\bar{X}_{ij.} - \bar{X}_{i..} \right)^2,$$

$$\text{df} = \sum_{i=1}^{3} \left(\beta_i - 1 \right) = \beta_1 + \beta_2 + \beta_3 - 3;$$

$$SS_A = \sum_{i=1}^{3} N_i \left(\bar{X}_{i..} - \bar{X} \right)^2, \quad \text{df} = 2.$$

As usual, each SS divided by its df yields the corresponding MS. Two of these, $MS_{B(A)}$ and MS_A, become the numerators of F ratios for testing the significances of the two main effects, B and A, respectively. (The B or teacher effect[2] is only of incidental interest; the A effect — that of the different initial reading programs — is of primary concern in this study.)

The numerators for the relevant F ratios are readily computed as indicated above. The denominator (or "error MS") appropriate for each F ratio is much

---

[2] In the body of the technical report, a more general term "classroom effect" is used instead of "teacher effect." It was felt that not only the teacher, but other factors such as instruction, peer group, etc. were unique for each classroom.

more difficult to obtain — especially when β_i (the number of B levels nested in A_i) varies from one A level to another, as it does in this study, as well as n_{ij} varying from one class to another. The general principle is as follows:

> Each error MS must be such that, under the corresponding null hypothesis [no teacher effect (i.e., B(A) effect) and no program effect (i.e. A effect) respectively] the numerator MS will have an expected value equal to the expected value of that error MS.

According to Kempthorne (1952, pp. 108-109), with one correction by the present author, the relevant expected values are as follows (where we abbreviate "within B" as "w – B"):

$$E\left(MS_{w-B}\right) = \sigma^2,$$

$$E\left(MS_{B(A)}\right) = \sigma^2 + P\ \sigma_b^2,$$

$$E\left(MS_A\right) = \sigma^2 + Q\ \sigma_b^2 + R\ \sigma_a^2,$$

where

$$P = \left[N - \sum_{i=1}^{3} \left(\frac{\sum_{j=1}^{\beta_i} n_{ij}^2}{N_i} \right) \right] \Bigg/ \sum_{i=1}^{3} (\beta_i - 1),$$

$$Q = \frac{1}{2} \left[\sum_{i=1}^{3} \left(\frac{\sum_{j=1}^{\beta_i} n_{ij}^2}{N_i} \right) - \frac{\sum_{i=1}^{3} \sum_{j=1}^{\beta_i} n_{ij}^2}{N} \right],$$

$$R = \frac{1}{2} \left(N - \frac{\sum_{i=1}^{3} N_i^2}{N} \right),$$

and

σ^2 is the variance due to sampling error (i.e., individual differences among pupils)

σ_a^2 is the variance due to program effects,

σ_b^2 is the variance due to teacher effects.

It is thus seen that the appropriate error MS for $MS_{B(A)}$ (i.e., the denominator of the F ratio for testing the teacher effect) is simply MS_{w-B} itself. For

$E(MS_{B(A)}) = \sigma^2 + P\,\sigma_b^2$ reduces to σ^2, which is $E(MS_{w-B})$,

when $\sigma_b^2 = 0$.

The appropriate error MS for MS_A [i.e., the denominator of the F ratio for testing the program effect (which is of prime interest)], on the other hand, must be constructed from a linear combination

$$L = C_1 (MS_{w-B}) + C_2 (MS_{B(A)})$$

in such a way that

$$E(L) = \sigma^2 + Q\,\sigma_b^2,$$

which is what $E(MS_A)$ reduces to when $\sigma_a^2 = 0$ (i.e., under the null hypothesis of "no program effect").

It may be verified that the above condition for L is satisfied if (and only if)

we take the following values for C_1 and C_2 :

$$C_1 = \frac{P - Q}{P}, \qquad C_2 = \frac{Q}{P},$$

where P and Q are as defined earlier. Thus, the appropriate error MS for MS_A is given by:

$$MS_{e(A)} = \left(\frac{P - Q}{P}\right) MS_{w-B} + \left(\frac{Q}{P}\right) MS_{B(A)} .$$

The only remaining question is that of the df for this error MS. There is apparently no universal consensus among statisticians on this question, but one widely accepted solution is that given by Mood (1950, pp. 334-348). Following this approach, the df for $MS_{e(A)}$ would be given by

$$df_{e(A)} = \frac{\left[C_1 \; MS_{w-B} + C_2 \; MS_{B(A)}\right]^2}{\frac{C_1^{\,2} \left(MS_{w-B}\right)^2}{N - (\beta_1 + \beta_2 + \beta_3)} + \frac{C_2^{\,2} \left(MS_{B(A)}\right)^2}{\beta_1 + \beta_2 + \beta_3 - 3}},$$

where C_1 and C_2 are as specified above.

This completes our description of the routine used in each of the 72 two-factor hierarchal design analyses used in this study. The actual computations were carried out at the University of Illinois Computer Center by Dr. Doyle Bishop, using a program which he wrote especially for these analyses.

APPENDIX B

MEAN SCORE PROFILES
FOR GRADES ONE, TWO, AND THREE

FIGURE 7

MEAN SCORES OF HIGH ABILITY PUPILS
ON THE STANFORD ACHIEVEMENT TEST
AT GRADE PLACEMENT 1.5

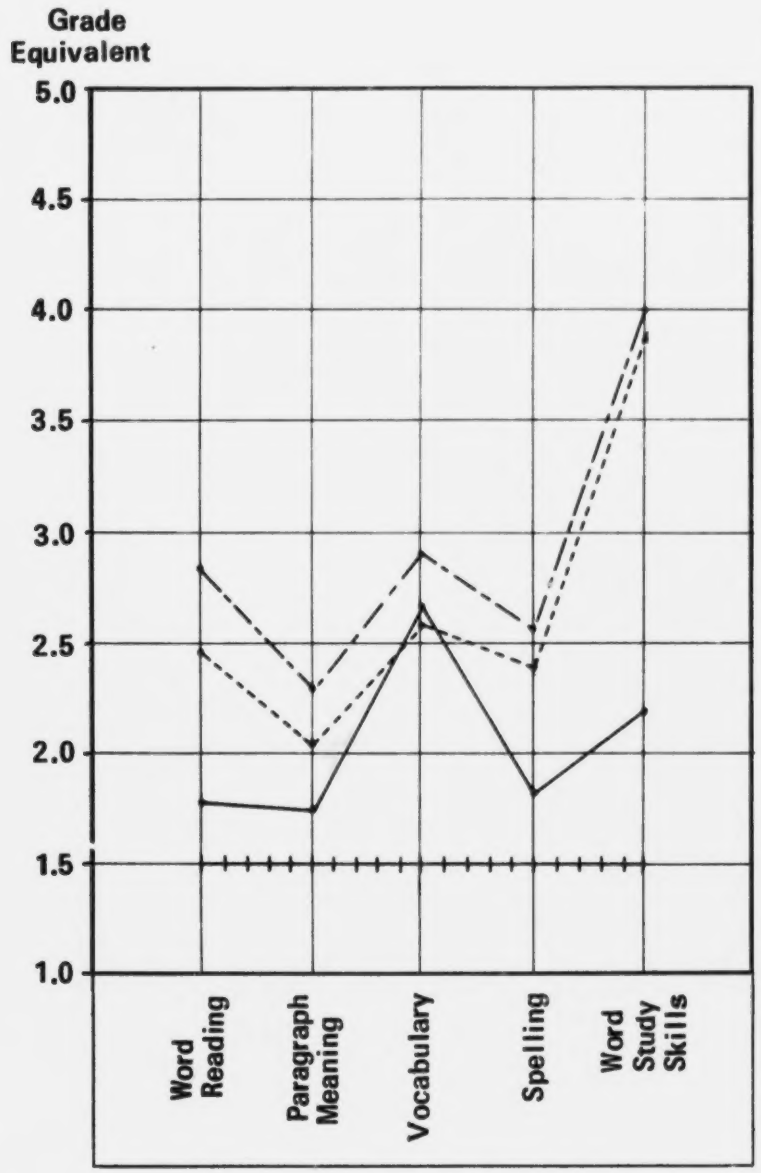


FIGURE 8

MEAN SCORES OF MIDDLE ABILITY PUPILS
ON THE STANFORD ACHIEVEMENT TEST
AT GRADE PLACEMENT 1.5

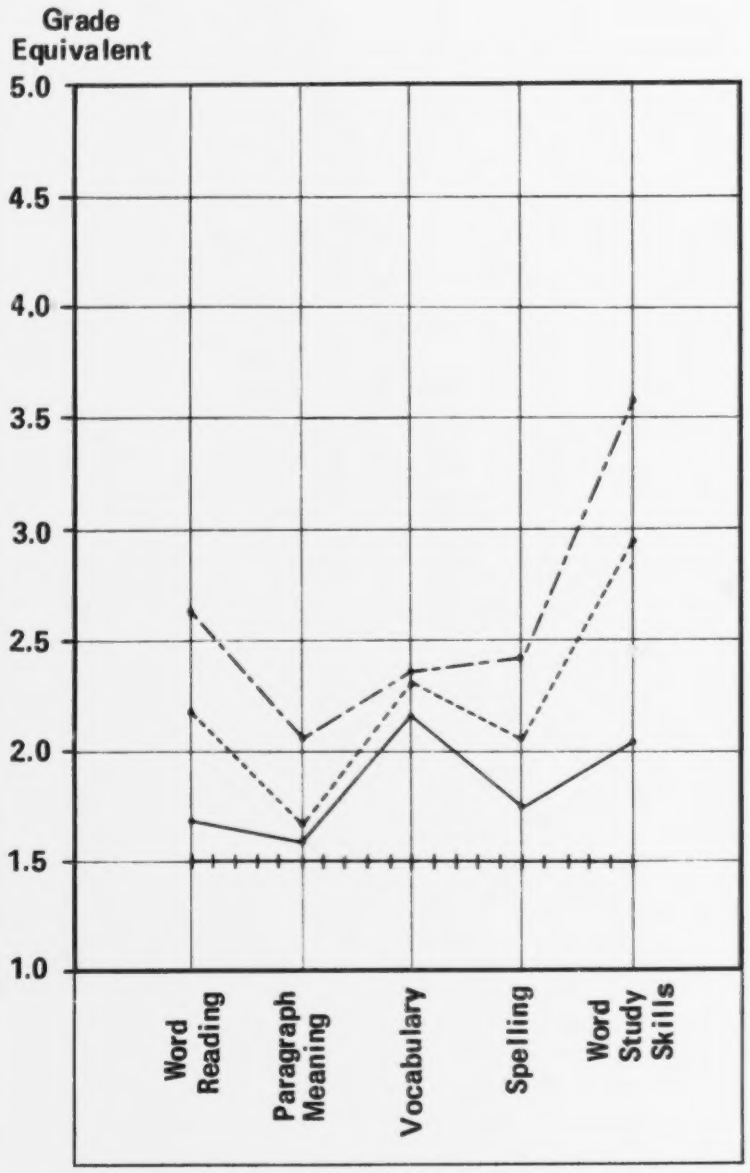
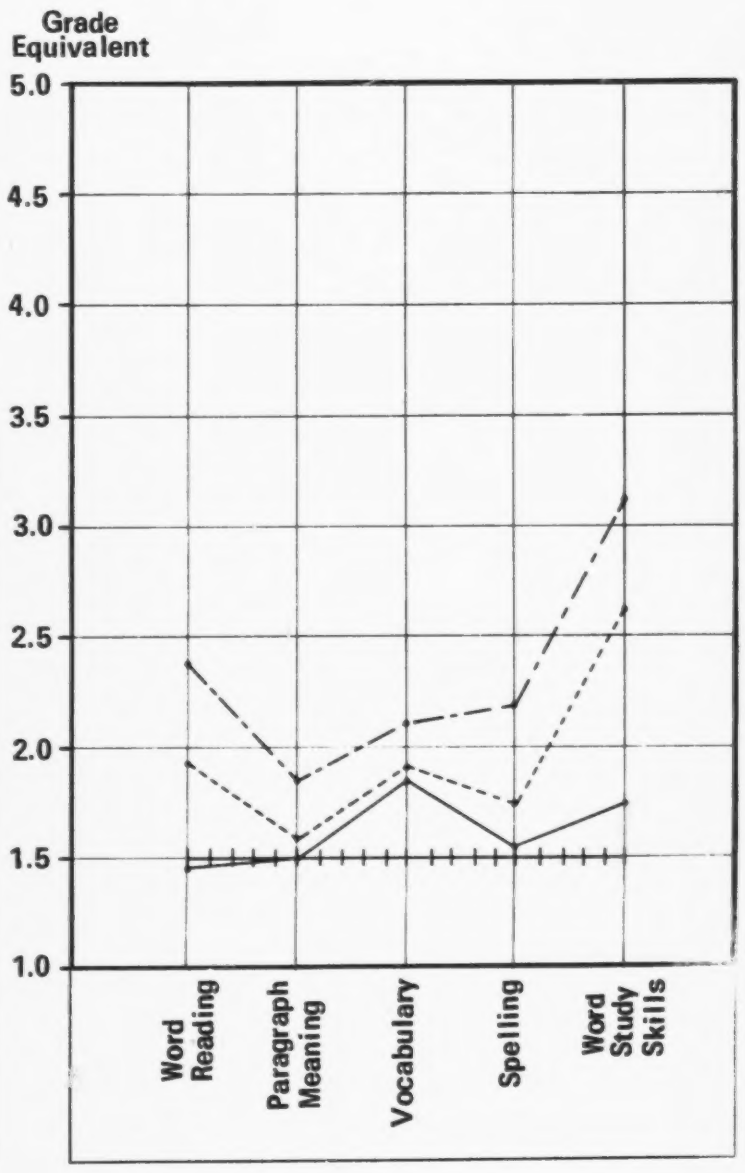


FIGURE 9

MEAN SCORES OF LOW ABILITY PUPILS
ON THE STANFORD ACHIEVEMENT TEST
AT GRADE PLACEMENT 1.5



i.t.a. 1 ---------- i.t.a. K —·—·—·—·—
T.O. ———————— Grade placement ++++

FIGURE 10

MEAN SCORES OF HIGH ABILITY PUPILS
ON THE STANFORD ACHIEVEMENT TEST
AT GRADE PLACEMENT 1.9

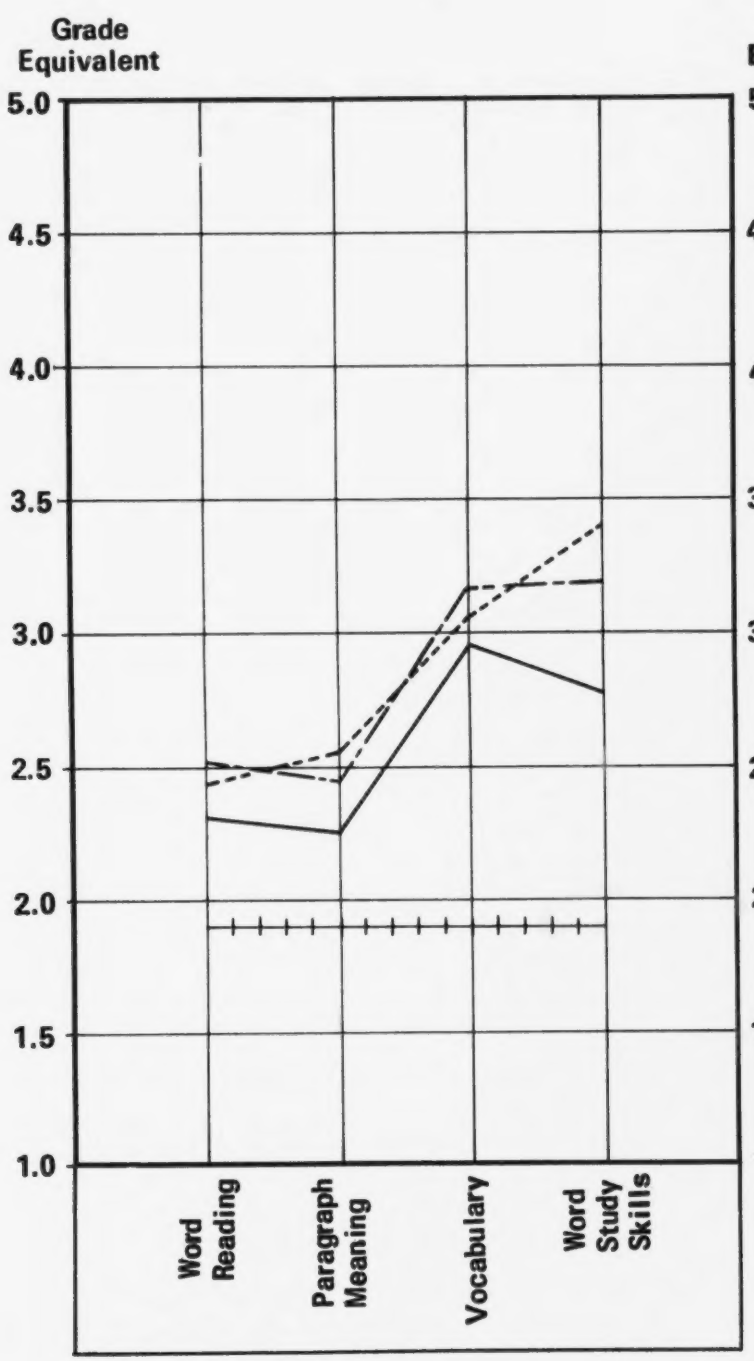


FIGURE 11

MEAN SCORES OF MIDDLE ABILITY PUPILS
ON THE STANFORD ACHIEVEMENT TEST
AT GRADE PLACEMENT 1.9

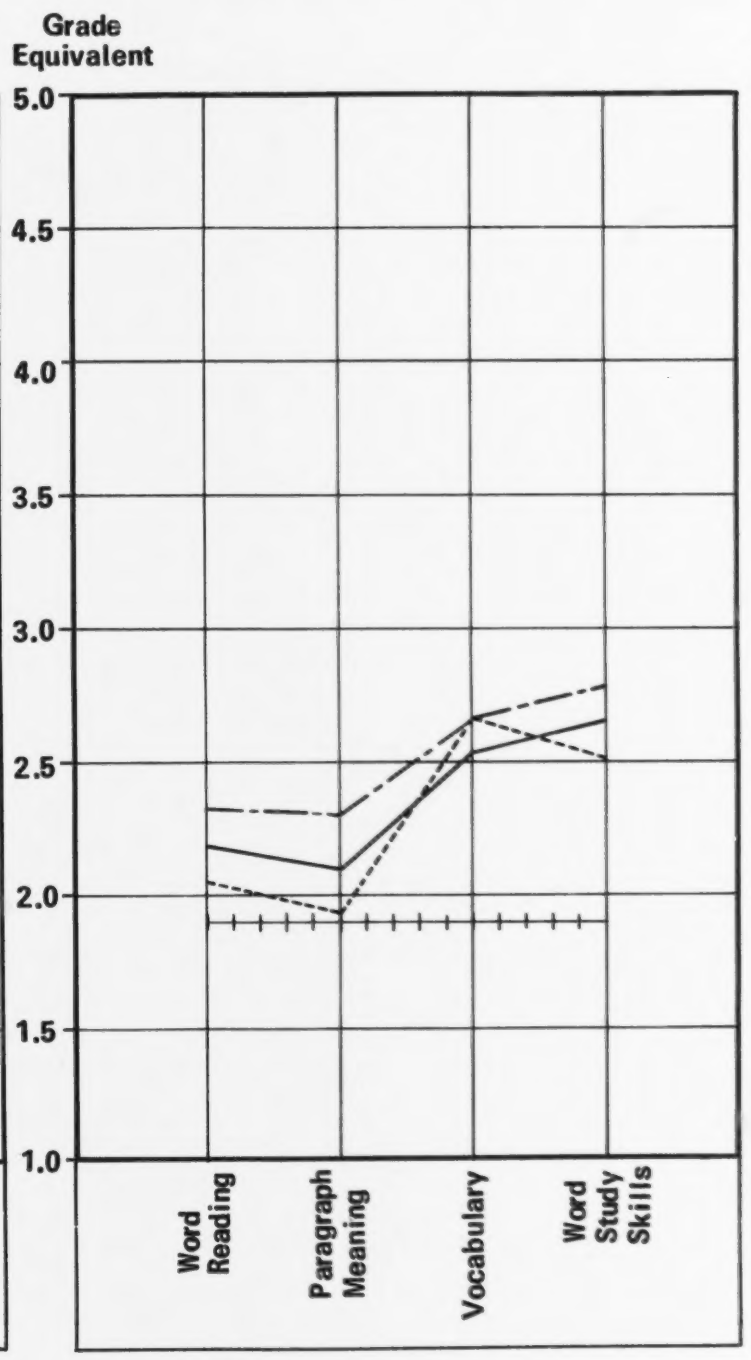
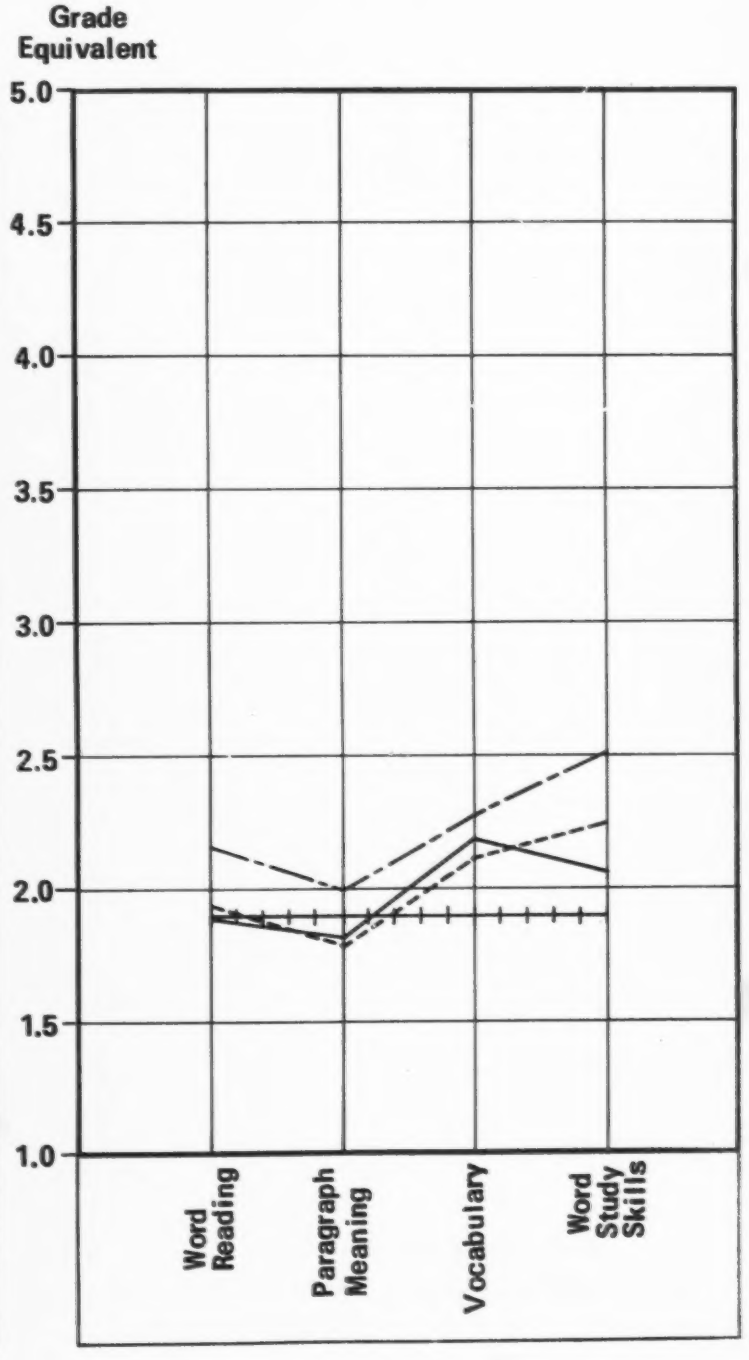


FIGURE 12

MEAN SCORES OF LOW ABILITY PUPILS
ON THE STANFORD ACHIEVEMENT TEST
AT GRADE PLACEMENT 1.9



i.t.a. 1 ---------- i.t.a. K —·—·—·—·—
T.O. ———————— Grade placement +++++++

FIGURE 13

MEAN SCORES OF HIGH ABILITY PUPILS
ON THE STANFORD ACHIEVEMENT TEST
AT GRADE PLACEMENT 2.8

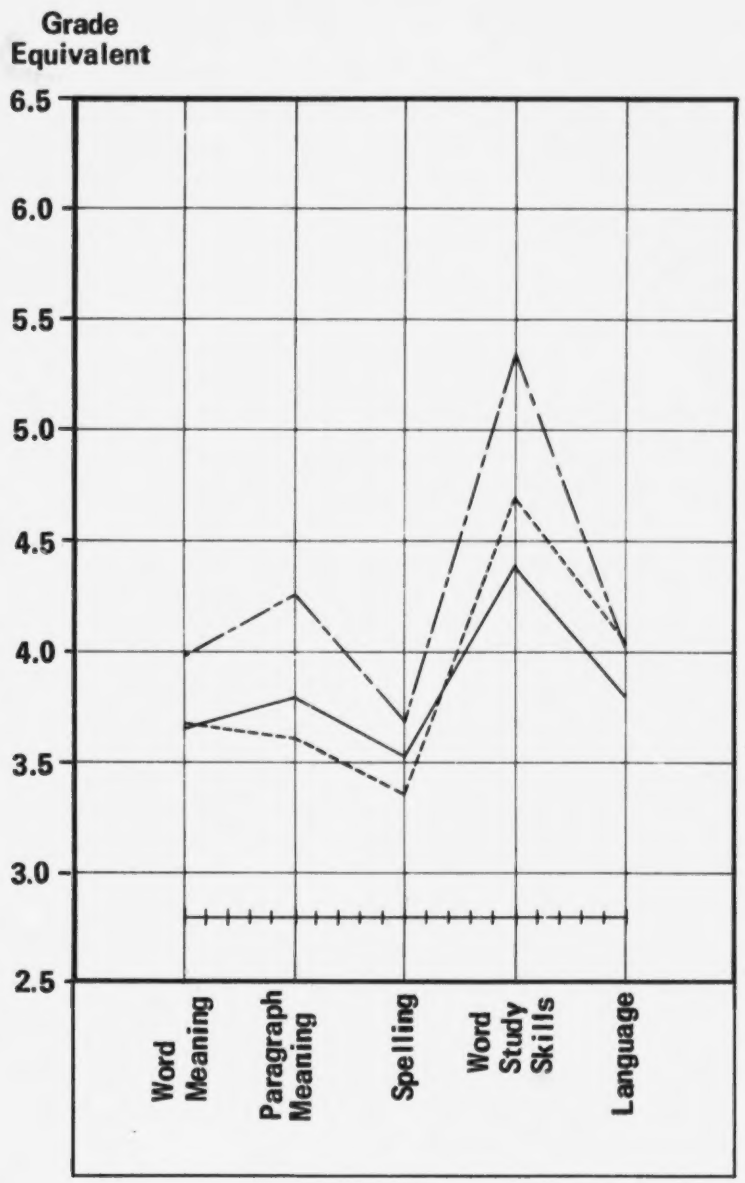


FIGURE 14

MEAN SCORES OF MIDDLE ABILITY PUPILS
ON THE STANFORD ACHIEVEMENT TEST
AT GRADE PLACEMENT 2.8

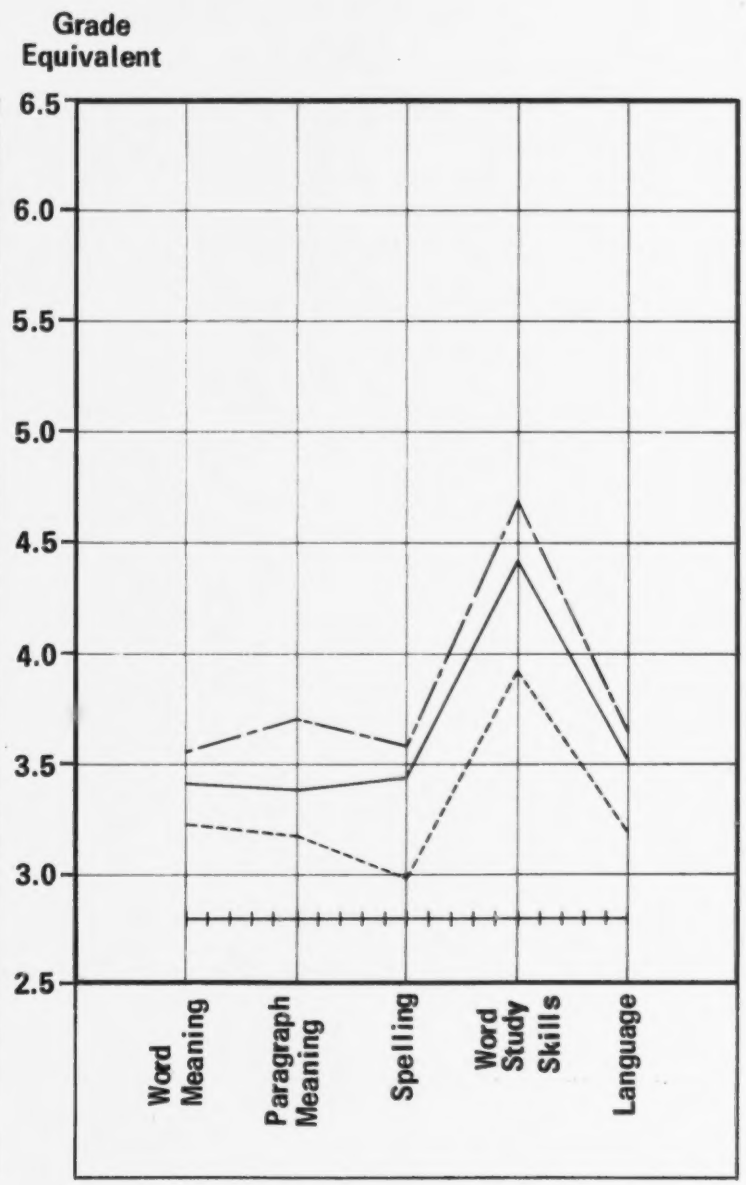
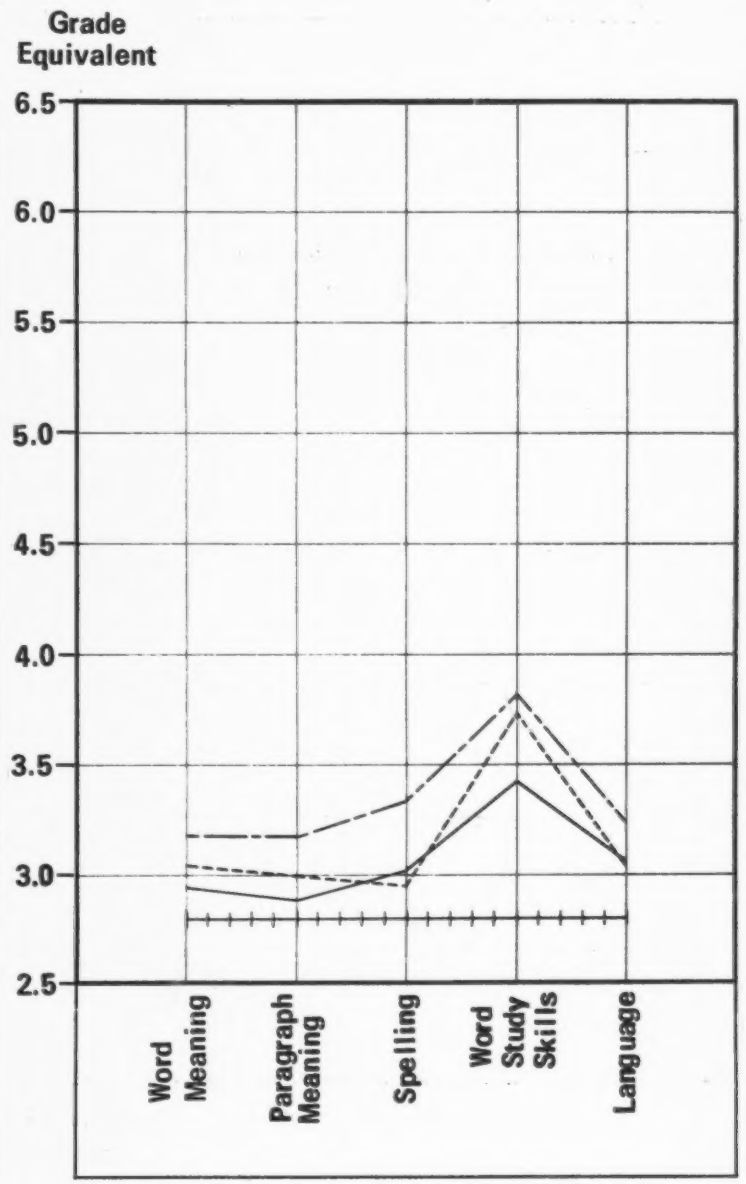


FIGURE 15

MEAN SCORES OF LOW ABILITY PUPILS
ON THE STANFORD ACHIEVEMENT TEST
AT GRADE PLACEMENT 2.8



i.t.a. 1 --------- i.t.a. K — - — - —
T.O. ———————— Grade placement ++++

FIGURE 16

MEAN SCORES OF HIGH ABILITY PUPILS ON THE STANFORD ACHIEVEMENT TEST AT GRADE PLACEMENT 3.9

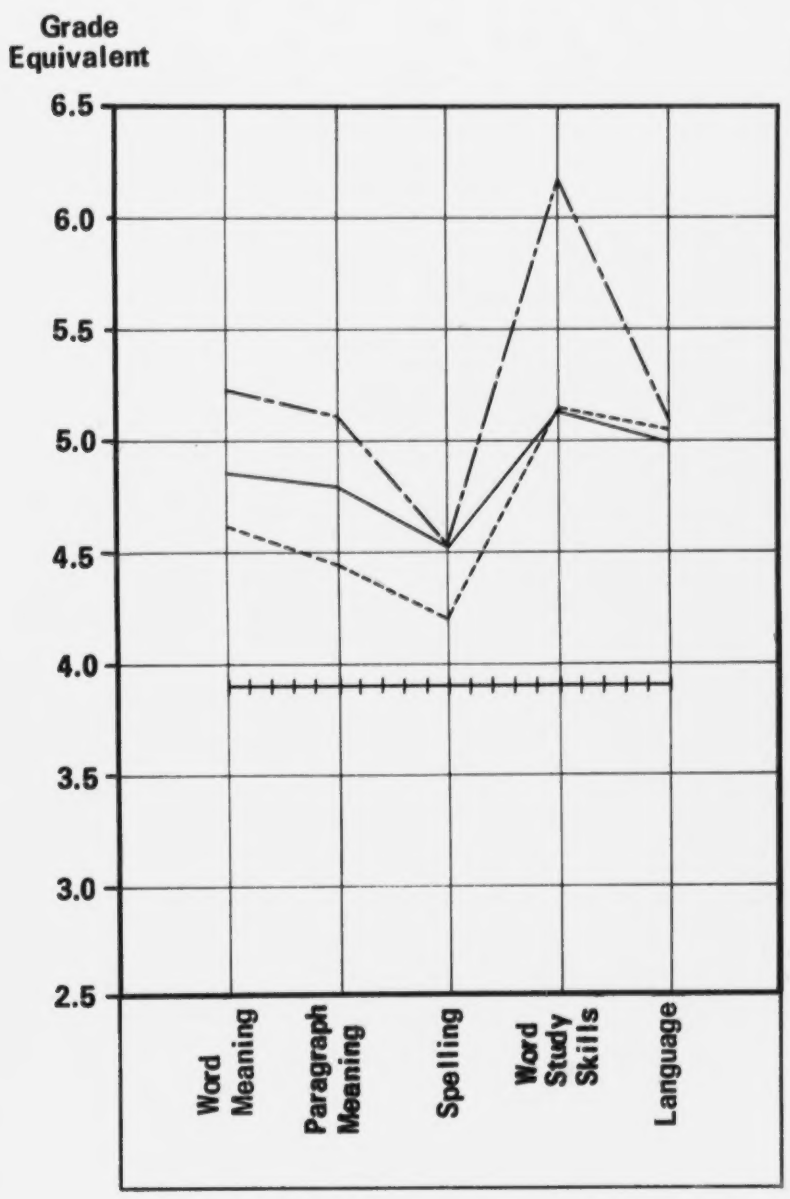


FIGURE 17

MEAN SCORES OF MIDDLE ABILITY PUPILS ON THE STANFORD ACHIEVEMENT TEST AT GRADE PLACEMENT 3.9

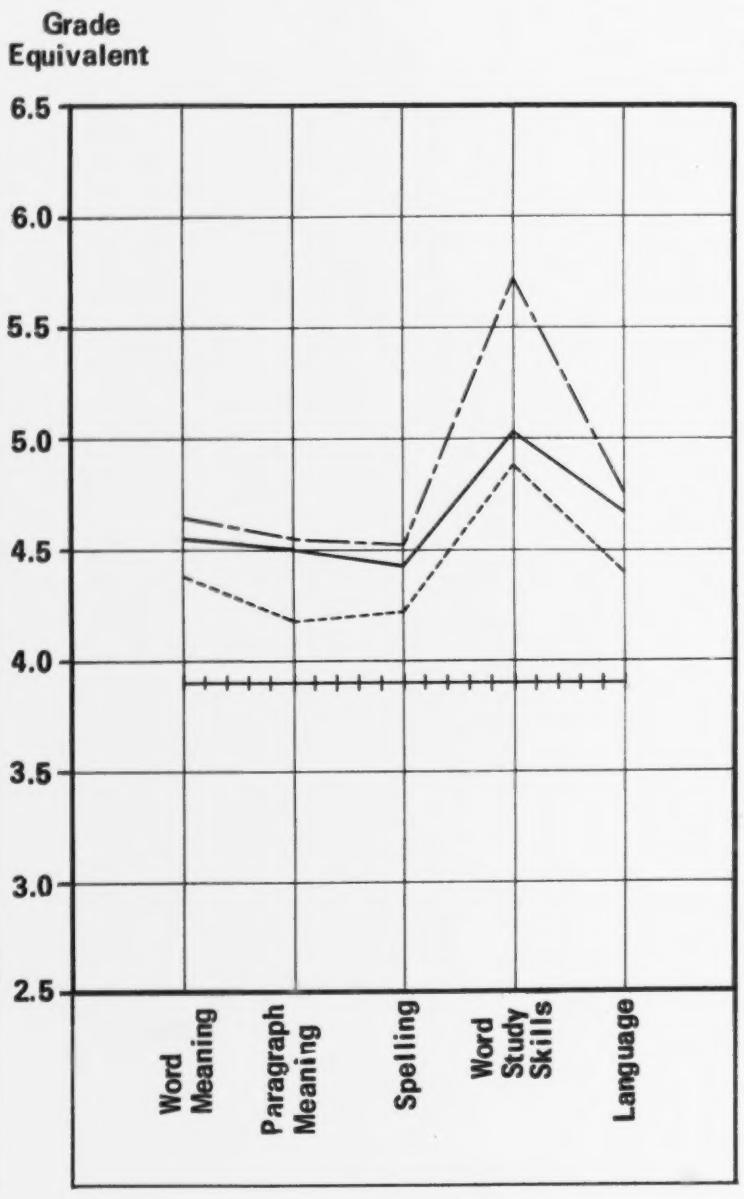
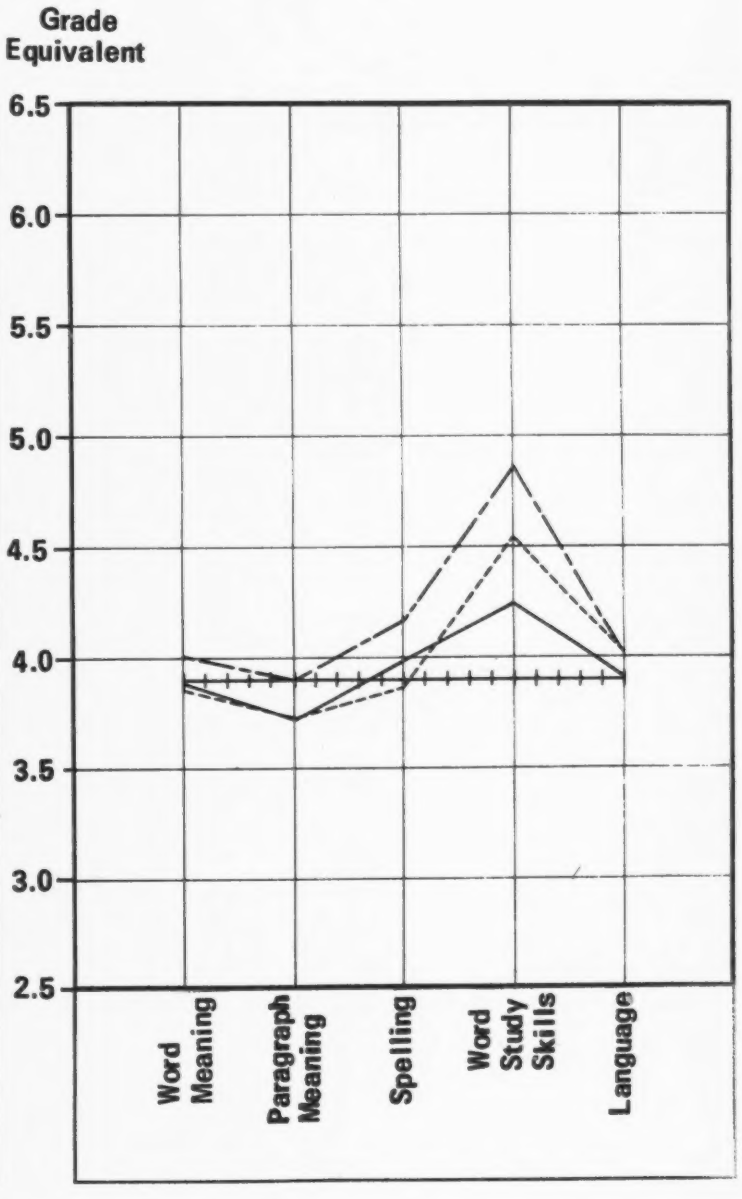


FIGURE 18

MEAN SCORES OF LOW ABILITY PUPILS ON THE STANFORD ACHIEVEMENT TEST AT GRADE PLACEMENT 3.9



i.t.a. 1 -------- i.t.a. K — — — — —
T.O. ———————— Grade placement |+|+|+|

FIGURE 19

MEAN TEACHER RATINGS
OF HIGH ABILITY PUPILS
AT GRADE PLACEMENT 3.9

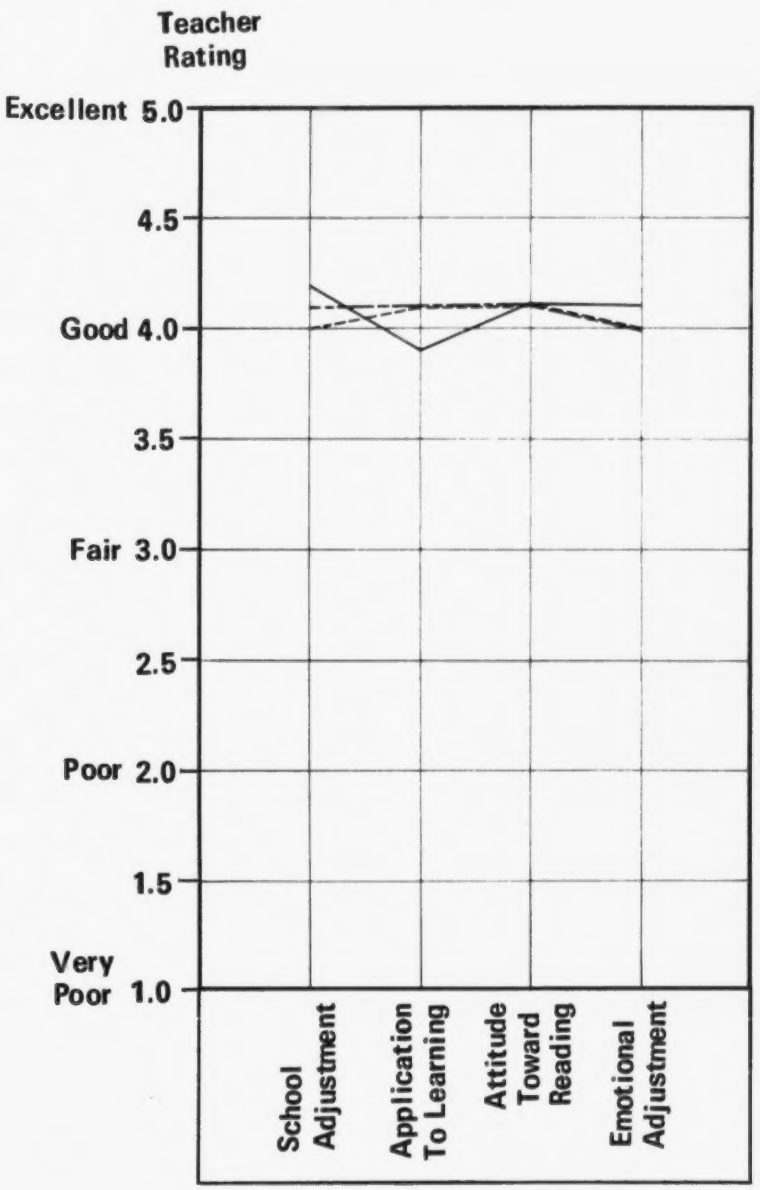


FIGURE 20

MEAN TEACHER RATINGS
OF MIDDLE ABILITY PUPILS
AT GRADE PLACEMENT 3.9

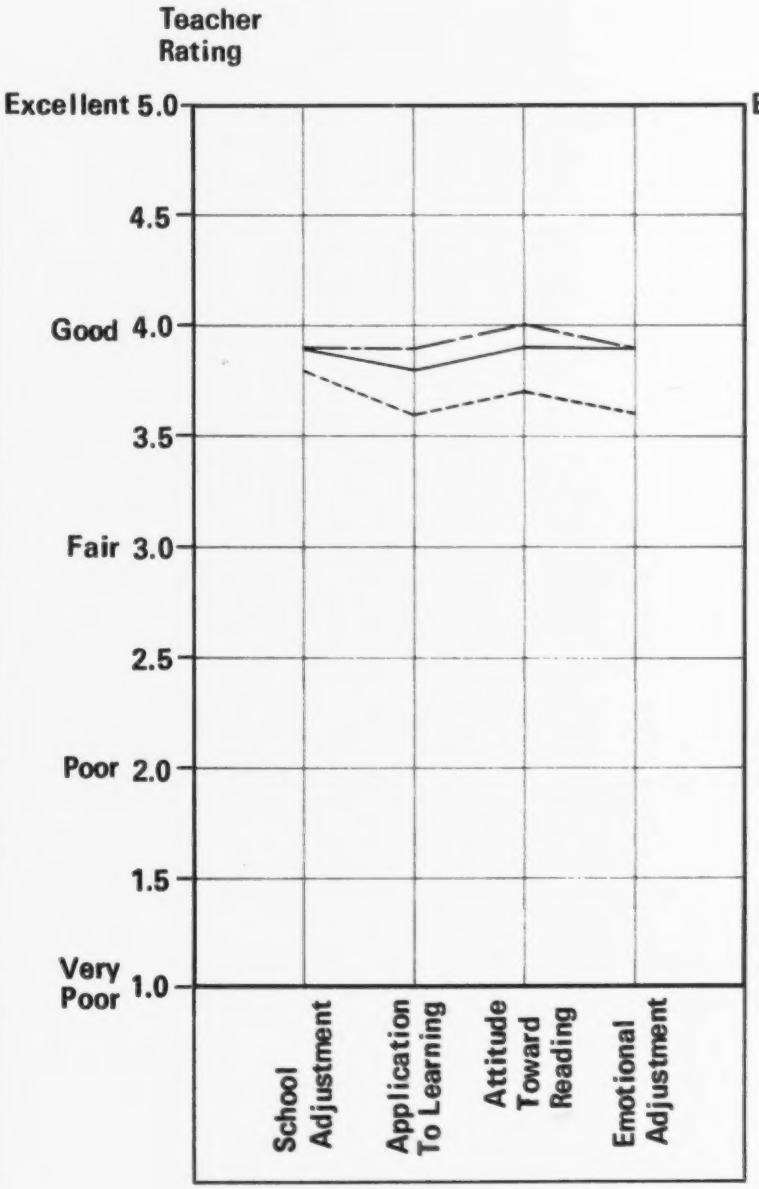
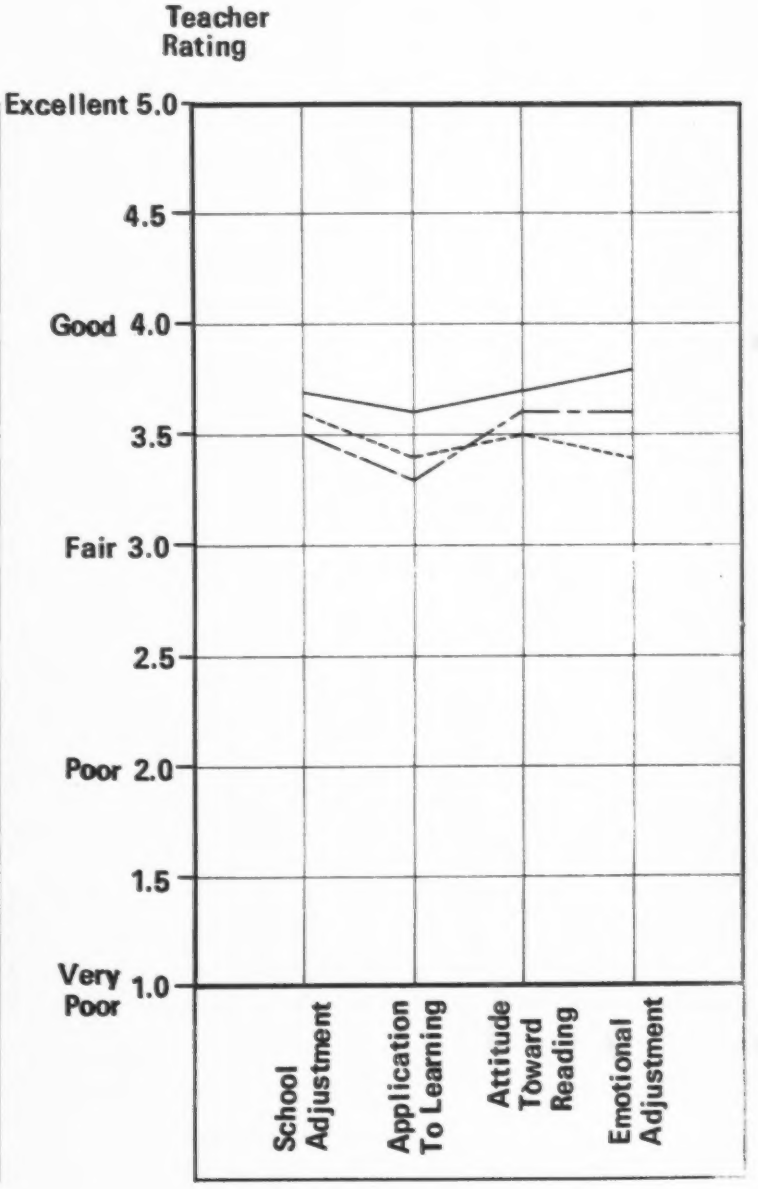


FIGURE 21

MEAN TEACHER RATINGS
OF LOW ABILITY PUPILS
AT GRADE PLACEMENT 3.9



i.t.a. 1 ------
T.O. ————————
i.t.a. K —— - ——

REFERENCES

Asher, J. W. Comment on "A comparison of i.t.a. and T.O. reading achievement when methodology is controlled." Elementary English, 1968, 45, 452-457, + 484.

Block, J. R. A critique of research with the initial teaching alphabet and some recommendations. The i.t.a. Foundation Report, 1966, 1 (2 & 3), 32-42.

Chall, Jeanne S. Learning to read: The great debate. New York: McGraw-Hill, 1967.

Chasnoff, R. E. Two alphabets. Elementary School Journal, 1967, 67, 257-264.

Chasnoff, R. E. Two alphabets: A follow-up. Elementary School Journal, 1968, 68, 251-257.

Fry, E. B. First grade reading instruction using diacritical marking system, initial teaching alphabet and basal reading system. The Reading Teacher, 1966, 19, 666-669.

Fry, E. B. First grade reading instruction using diacritical marking system, initial teaching alphabet and basal reading system — extended to second grade. The Reading Teacher, 1967, 20, 687-693. (a)

Fry, E. B. i.t.a.: A look at the research data. Education, 1967, 87, 549-553. (b)

Fry, E. B. Comparison of beginning reading with i.t.a., DMS, and T.O. after three years. The Reading Teacher, 1969, 22, 357-362.

Gillooly, W. B. The use of i.t.a. in special education: A critical review. Journal of Special Education, 1967, 1, 127-134.

Gillooly, W. B. The effectiveness of i.t.a. in reading instruction. Paper read at International Reading Association, Boston, April 27, 1968.

Hahn, H. T. Three approaches to beginning reading instruction — i.t.a., language arts, and basic readers. The Reading Teacher, 1966, 19, 590-594.

Hahn, H. T. Three approaches to beginning reading instruction — i.t.a., language experience and basic readers — extended into second grade. The Reading Teacher, 1967, 20, 711-715.

Hayes, R. B. i.t.a. and three other approaches to reading in first grade. The Reading Teacher, 1966, 19, 627-630.

Hayes, R. B., & Wuest, R. C. i.t.a. and three other approaches to reading in first grade — extended into second grade. The Reading Teacher, 1967, 20, 694-697, + 703.

Hayes, R. B., & Wuest, R. C. A three year look at i.t.a., Lippincott, Phonics and Word Power, and Scott, Foresman. The Reading Teacher, 1969, 22, 363-370.

Kempthorne, O. The design and analysis of experiments. New York: Wiley, 1952.

Mazurkiewicz, A. J. i.t.a. and T.O. reading achievement when methodology is controlled. The Reading Teacher, 1966, 19, 606-610.

Mazurkiewicz, A. J. i.t.a. and T.O. reading achievement when methodology is controlled — extended into second grade. The Reading Teacher, 1967, 20, 726-729.

Mood, A. H. Introduction to the theory of statistics. New York: McGraw-Hill, 1950.

Tanyzer, H. J., & Alpert, H. Three different basal reading systems and first grade reading achievement. The Reading Teacher, 1966, 19, 636-642.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

Council Staff*

Dr. George H. Baird	Executive Director
Mr. L. Romanos	Associate Director
Dr. Alden H. Blankenship	Associate Director
Mr. Carl Dautel	Business Manager
Mr. Ben W. Wood	Consultant
Dr. Harry S. Broudy	Consultant
Dr. Ralph H. Ojemann	Director, Child and Educational Psychology-Preventive Psychiatry
Dr. Wai-Ching Ho	Director, Evaluation and Testing
Mr. Raymond English	Director, Social Science
Mr. John F. Mehegan	Director, Mathematics
Mr. Bernard J. McCabe	Director, English and Language Arts
Dr. James Joseph Gallagher	Director, Science
Mr. Lester Smith	Director, Health and Physical Education
Mr. John W. Dye	Liaison Director
Mrs. Eleanor G. Zirzow	Coordinator, Reading and i.t.a.
Mr. Michael Constant	Coordinator, Humanities
Mlle. Marthe G. Garnier	Coordinator, Project French
Mr. Rupert Beckstett	Coordinator, Occupational Education
Mr. Robert L. Pegues, Jr.	Coordinator, Urban Education
Mr. Lewis F. Mayer	Coordinator, Administrative Services
Mr. Joseph E. Bechard	Coordinator, Staff Utilization
Mr. Richard Parker	Coordinator, In-Service Education
Mr. Robert A. Murphy	Coordinator, Multimedia Development Center
Mr. Theodore N. Ferris, Jr.	Editor, Central Office
Mr. Elmer F. Bowhall	Manager, Data Processing

Consultants

Evaluation and Testing: Dr. John B. Carroll • Dr. J. Thomas Hastings • Dr. Maurice M. Tatsuoka • Dr. Ralph W. Tyler **Humanities:** Dr. Harry S. Broudy **In-Service Education & Staff Utilization:** Dr. J. Lloyd Trump **Mathematics:** Dr. Max Beberman • Dr. Charles Buck • Mr. George S. Cunningham • Dr. Jack E. Forbes • Dr. George Polya **Psychology:** Dr. Leon J. Saul • Dr. Silas Warner **Science:** Dr. Garland E. Allen • Dr. William W. Hambleton • Dr. Paul DeHart Hurd • Dr. L. Carroll King • Dr. Addison E. Lee • Dr. Joseph D. Novak • Dr. Clifford E. Swartz • Dr. Fletcher G. Watson **Social Science:** Dr. Phillip Bacon • Dr. Harold F. Clark • Dr. Russell Kirk • Dr. William H. McNeill • Dr. James M. McPherson

Evaluation and Testing Department

Dr. Wai-Ching Ho	Director
Dr. Thomas E. Bibler	Administrative Associate
Dr. Charles F. Eiszler, Jr.	Research Associate
Dr. Edwin F. Sause	Research Associate
Mrs. Susan Detienne	Research Assistant
Miss Sarita Schrock	Research Assistant
Mr. Michael Swanson	Research Fellow

*** March 1970**